

Mon Jun 17 16:11:22 2002

us-09-761-636a-14.open.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:40 ; Search time 75.98 Seconds
(without alignments)
2.893 Million cell updates/sec

Title: US-09-761-636a-14
Perfect score: 52
Sequence: 1 C1SVPLVPC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTUTS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	41	78.8	342	1 US-08-499-215-5	Sequence 5, Appl
2	35	67.3	1440	4 US-09-357-251-37	Sequence 37, Appl
3	34	65.4	215	1 US-08-107-684B-11	Sequence 11, Appl
4	34	65.4	1876	2 US-08-609-049A-12	Sequence 12, Appl
5	34	65.4	1876	2 US-08-609-049A-28	Sequence 28, Appl
6	34	65.4	1876	4 US-09-170-996-12	Sequence 12, Appl
7	34	65.4	1876	4 US-09-170-996-28	Sequence 28, Appl
8	33	63.5	20	4 US-09-461-697-169	Sequence 169, App
9	33	63.5	32	4 US-09-461-697-167	Sequence 167, App
10	33	63.5	1253	1 US-07-920-281C-3	Sequence 3, Appl
11	33	63.5	1253	4 US-08-466-277-3	Sequence 3, Appl
12	32	61.5	185	1 US-07-988-273-2	Sequence 2, Appl
13	32	61.5	185	3 US-08-848-810-25	Sequence 25, Appl
14	32	61.5	185	3 US-09-164-193-21	Sequence 21, Appl
15	32	61.5	185	4 PCT-US93-12019-2	Sequence 2, Appl
16	32	61.5	218	5 US-08-985-526-1	Sequence 1, Appl
17	32	61.5	239	5 PCT-US93-01652-1	Sequence 1, Appl
18	32	61.5	305	4 US-09-248-588-13	Sequence 13, Appl
19	32	61.5	441	3 US-08-985-526-3	Sequence 3, Appl
20	32	61.5	663	4 US-08-959-004-5	Sequence 5, Appl
21	32	61.5	1170	1 US-08-313-288B-20	Sequence 20, Appl
22	32	61.5	22	4 US-09-227-357-517	Sequence 517, App
23	31	59.6	144	2 US-08-950-168-3	Sequence 3, Appl
24	31	59.6	144	4 US-09-365-705-3	Sequence 3, Appl
25	31	59.6	159	3 US-08-928-361B-9	Sequence 9, Appl
26	31	59.6	268	2 US-08-824-874-1	Sequence 1, Appl
27	31	59.6	268	4 US-09-210-084-1	Sequence 1, Appl

28	31	59.6	496	4 US-09-171-969-2	Sequence 2, Appl
29	31	59.6	597	6 5171850-2	Patent No. 5171850
30	31	59.6	603	3 US-08-482-677-8	Sequence 8, Appl
31	31	59.6	604	2 US-08-635-137-2	Sequence 2, Appl
32	31	59.6	604	4 US-09-136-981-2	Sequence 19, Appl
33	31	59.6	871	4 US-09-245-041-19	Sequence 50, Appl
34	31	59.6	1837	4 US-08-928-361B-5	Sequence 230, App
35	30	57.7	39	4 US-09-227-357-230	Sequence 17, Appl
36	30	57.7	63	4 US-09-188-930-132	Patent No. 5208218
37	30	57.7	144	6 5208218-1	Sequence 17, Appl
38	30	57.7	382	1 US-08-318-947A-17	Sequence 17, Appl
39	30	57.7	382	2 US-08-795-303-17	Patent No. 5268463
40	30	57.7	426	6 5268463-8	Patent No. 5432081
41	30	57.7	428	6 5432081-9	Patent No. 5432081
42	30	57.7	456	6 5432081-7	Sequence 6, Appl
43	30	57.7	457	2 US-08-882-704A-6	Patent No. 5268463
44	30	57.7	457	6 5268463-7	Sequence 89, Appl
45	30	57.7	510	4 US-09-522-217-89	

ALIGNMENTS

RESULT 1
US-08-499-215-5
Sequence 5, Application US/08499215
Patent No. 5612204
GENERAL INFORMATION:
APPLICANT: Saeki, Hisashi
TITLE OF INVENTION: BIOLOGICAL DEGRADATIVE
TITLE OF INVENTION: TREATMENT OF CHLORINE-SUBSTITUTED ETHYLENE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue
CITY: N.W.
STATE: Washington D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentPerfect version 5.1
SOFTWARE: 6 WordPerfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/499,215
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-179689
FILING DATE: 08-JUL-1994
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-499-215-5

Query Match 78.8%, Score 41: DB 1: Length 342:
Best Local Similarity 77.8%, Pred. No. 14:
Matches 7: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

QY 1 C1SVPLVPC 9
Db 75 CASVPLVPC 83

RESULT 2
US-09-357-251-37

TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,049A
FILING DATE: 29-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1876 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-28

Query Match 65.4%; Score 34; DB 2; Length 1876;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTSVPLVPC 9
DB 848 CWSKPIVVC 856

RESULT 6
US-09-170-996-12
Sequence 12, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Molz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1876 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-12

Query Match 65.4%; Score 34; DB 4; Length 1876;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CTSVPLVPC 9
DB 848 CWSKPIVVC 856

RESULT 7
US-09-170-996-28
Sequence 28, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Molz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1876 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-28

Query Match 65.4%; Score 34; DB 4; Length 1876;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
1:1:1:1
DB 848 CVSKPIVVC 856

RESULT 8

US-09-461-697-169
; Sequence 169, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: LO, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Putnam, Kasturi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-169

Query Match
Best Local Similarity 63.5%; Score 33; DB 4; Length 20;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 C1SVPLVPC 9
1:1:1:1
DB 4 CSSIPVVC 12

RESULT 9

US-09-461-697-167
; Sequence 167, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: LO, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Putnam, Kasturi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-167

Query Match
Best Local Similarity 63.5%; Score 33; DB 4; Length 32;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
1:1:1:1
DB 16 CSSIPVVC 24

RESULT 10

US-07-920-281C-3
; Sequence 3, Application US/07920281C
; Patent No. 5739026
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; APPLICANT: Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; TITLE OF INVENTION: Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/920,281C
; FILING DATE: 13-AUG-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 828-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-920-281C-3

Query Match
Best Local Similarity 63.5%; Score 33; DB 1; Length 1253;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 C1SVPLVPC 9
1:1:1:1
DB 285 CFOPPCVPC 293

RESULT 11

US-08-466-277-3
; Sequence 3, Application US/08466277
; Patent No. 6190666
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; APPLICANT: Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; TITLE OF INVENTION: Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA

ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,277
FILING DATE: 06-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/920,281
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-466-277-3

Query Match 63.5%; Score 33; DB 4; Length 1253;
Best Local Similarity 55.6%; Pred. No. 9.8e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
1 1 1 1 1 1
DB 285 C1SVPLVPC 293

RESULT 12
US-07-988-273-2
Sequence 2, Application US/07988273
Patent No. 5512434
GENERAL INFORMATION:
APPLICANT: AARONSON, Stuart A.
APPLICANT: BOTARO, Donald P.
APPLICANT: ISHIBASHI, Toshio
APPLICANT: MIKI, Toru
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,273
FILING DATE: 19921214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/182 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-988-273-2

Query Match 61.5%; Score 32; DB 1; Length 185;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
1 1 1 1 1 1
DB 22 C1SVPLVPC 30

RESULT 13
US-08-848-810-25
Sequence 25, Application US/08848810
Patent No. 6074851
GENERAL INFORMATION:
APPLICANT: Deibel Jr., M. R.
APPLICANT: Yem, A. W.
APPLICANT: Wilson, C. L.
TITLE OF INVENTION: Catalytic Macro Molecules Having DCD25B
TITLE OF INVENTION: Like Activity
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pharmacia & Upjohn Company
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/848,810
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Woolton, Thomas A.
REGISTRATION NUMBER: 35,004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-833-6897
TELEFAX: 616-833-7914
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-848-810-25

Query Match 61.5%; Score 32; DB 3; Length 185;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
1 1 1 1 1 1
DB 22 C1SVPLVPC 30

RESULT 14
US-09-164-193-21
; Sequence 21, Application US/09164193C
; Patent No. 6258582
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSAPTP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
; FILE REFERENCE: MNT-051
; CURRENT APPLICATION NUMBER: US/09/164.193C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-164-193-21

Query Match 61.5%; Score 32; DB 4; Length 185;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
| | | | |
Db 22 C1SLPSQPC 30

RESULT 15
PCT-US93-12019-2
; Sequence 2, Application PC/TUS9312019
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12019
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,273
; FILING DATE: 14-DEC-1992
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-12019-2

Query Match 61.5%; Score 32; DB 5; Length 185;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
| | | | |
Db 22 C1SLPSQPC 30

Search completed: June 17, 2002, 15:56:40
Job time: 89 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:25 ; Search time 44.84 Seconds
(without alignments)
6.908 Million cell updates/sec

Title: US-09-761-636A-6
Perfect score: 46
Sequence: 1 CNEESLIC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	354	1	VEGD_HUMAN
2	40	87.0	326	1	VEGD_RAT
3	36	78.3	133	1	VEGH_ORF2
4	36	78.3	146	1	VEGA_SHEEP
5	36	78.3	164	1	VEGA_CAVO
6	36	78.3	190	1	VEGA_BOVIN
7	36	78.3	358	1	VEGD_MOUSE
8	35	76.1	495	1	FASC_STRPU
9	35	76.1	4427	1	PKSL_BACSU
10	34	73.9	500	1	CPJ1_RABIT
11	34	73.9	501	1	CPJ6_MOUSE
12	34	73.9	502	1	CPJ3_RAT
13	34	73.9	1722	1	RBB2_HUMAN
14	34	71.7	214	1	VEGA_MOUSE
15	33	71.7	214	1	VEGA_RAT
16	33	71.7	508	1	CP77_HORSE
17	33	71.7	551	1	RM27_HUMAN
18	33	71.7	551	1	RM27_MOUSE
19	33	71.7	1786	1	LMB1_HUMAN
20	32	69.6	190	1	VEGA_HORSE
21	32	69.6	190	1	VEGA_PIG
22	32	69.6	214	1	VEGA_CANFA
23	32	69.6	376	1	VEGA_HUMAN
24	32	69.6	429	1	DXR_THEMA
25	32	69.6	432	1	MDHP_MATE
26	32	69.6	432	1	MDHP_SORBI
27	32	69.6	432	1	MDHP_SORBI
28	32	69.6	435	1	MDHP_SPIOL
29	32	69.6	437	1	MDHP_MEDSA
30	32	69.6	441	1	MDHP_MESGR
31	32	69.6	453	1	MDHP_PEA
32	31	67.4	51	1	MDHP_FLARI
33	31	67.4	51	1	YDAF_ECOLI

34	31	67.4	82	1	VG13_HSV11
35	31	67.4	171	1	NIDM_HUMAN
36	31	67.4	218	1	PUR0_METTH
37	31	67.4	231	1	TRMD_MYCCE
38	31	67.4	251	1	Y206_CHLPP
39	31	67.4	304	1	HEM3_AQUAE
40	31	67.4	353	1	HEM2_BRAVA
41	31	67.4	393	1	ACKA_MYCCA
42	31	67.4	541	1	YKCS_CAEEL
43	31	67.4	740	1	GNF5_RAT
44	31	67.4	741	1	GNF5_HUMAN
45	31	67.4	822	1	VG1B_HSV7J

ALIGNMENTS

RESULT	ID	VEGD_HUMAN	STANDARD	PRT	354 AA.
AC	043915				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).				
GN	FIGF OR VEGFD.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lung;				
RX	MEDLINE=97349118; PubMed=9205122;				
RA	Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;				
RT	"Molecular cloning of a novel vascular endothelial growth factor, VEGF-D."				
RL	Genomics 42:483-488(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lung;				
RX	MEDLINE=98140120; PubMed=9479493;				
RA	Rocchigiani M., Lestingi M., Lucidi A., Orlandini M., Franco B., Rossi E., Ballabio A., Zuffardi O., Oliviero S.;				
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1 between the FIGA and the GRPR genes."				
RL	Genomics 47:207-216(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98118549; PubMed=9435229;				
RA	Achen M.G., Jeltsch M., Kuk E., Maekinen T., Vitall A., Wilks A.F., Alitalo K., Stackel S.A.;				
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4)."				
RL	Proc. Natl. Acad. Sci. U.S.A. 95:348-353(1998).				
RN	[4]				
RP	PROCESSING, AND SEQUENCE OF 89-94; 100-105 AND 206-213.				
RX	MEDLINE=20011413; PubMed=10542248;				
RA	Stacker S.A., Stenvers K.L., Caesar C., Vitall A., Domagala T., Niche E.C., Roufail S., Simpson R.J., Moritz R., Karpenen T., Alitalo K., Achen M.G.;				
RT	"Biosynthesis of vascular endothelial growth factor-D involves proteolytic processing which generates non-covalent homodimers."				
RL	J. Biol. Chem. 274:32127-32136(1999).				
CC	-1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.				
CC	-1- SUBUNIT: Homodimer; non-covalent and antiparallel.				

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Highly expressed in lung, heart, small
 CC intestine and fetal lung, and at lower levels in skeletal muscle,
 CC colon, and pancreas.
 CC -1- Ptm: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
 CC linked by disulfide bonds before secretion. The fully processed
 CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 CC bound by non-covalent interactions.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D89630; BAA24264.1; -
 CC EMBL: Y12863; CA73370.1; -
 CC EMBL: Y12864; CA73371.1; -
 CC EMBL: Y12865; CA73371.1; JOINED.
 CC EMBL: Y12866; CA73371.1; JOINED.
 CC EMBL: Y12867; CA73371.1; JOINED.
 CC EMBL: Y12868; CA73371.1; JOINED.
 CC EMBL: Y12869; CA73371.1; JOINED.
 CC EMBL: Y12870; CA73371.1; JOINED.
 CC EMBL: AJ000185; CA03942.1; -
 CC MIM: 300091; -
 CC HSSP: P15692; IVP.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF; 1.
 CC ProDom: PD001629; PDGF; 1.
 CC SMART: SM00141; PDGF; 1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS00278; PDGF_2; 1.
 CC Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
 CC Cleavage on pair of basic residues; Multigene family.
 CC SIGNAL 1 21
 CC PROPEP 22 88
 CC CHAIN 89 205
 CC PROPEP 206 354
 CC DOMAIN 222 318
 CC REPEAT 222 237
 CC REPEAT 258 273
 CC REPEAT 277 293
 CC REPEAT 301 318
 CC DISULFID 111 153
 CC DISULFID 142 189
 CC DISULFID 146 191
 CC DISULFID 136 136
 CC DISULFID 145 145
 CC CARBOHYD 155 155
 CC CARBOHYD 185 185
 CC CARBOHYD 287 287
 CC SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;
 Query Match 100.0%; Score 46; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. NO. 0.07; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CNEESLIC 8
 Db 146 CNEESLIC 153
 RESULT 2
 VEGD_RAT
 ID VEGD_RAT STANDARD; PRT; 326 AA.

AC 035251;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
 DE growth factor) (FIGF).
 GN FIGF OR VEGFD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN=Sprague-Dawley;
 RA Yamada Y., Hirata Y., Nezu J., Shitane M.;
 RL Submitted (Jul-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
 CC and endothelial cell growth, stimulating their proliferation and
 CC migration and also has effects on the permeability of blood
 CC vessels. May function in the formation of the venous and lymphatic
 CC vascular systems during embryogenesis, and also in the maintenance
 CC of differentiated lymphatic endothelium in adults. Binds and
 CC activates VEGFR-3 (Flt4) receptor (By similarity).
 CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- Ptm: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
 CC linked by disulfide bonds before secretion. The fully processed
 CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 CC bound by non-covalent interactions (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF014827; AAB6557.1; -
 CC HSSP: P15692; IVP.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF; 1.
 CC ProDom: PD001629; PDGF; 1.
 CC SMART: SM00141; PDGF; 1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS00278; PDGF_2; 1.
 CC Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
 CC Cleavage on pair of basic residues; Multigene family.
 CC SIGNAL 1 21
 CC PROPEP 22 93
 CC CHAIN 94 210
 CC PROPEP 211 326
 CC DOMAIN 227 317
 CC REPEAT 227 242
 CC REPEAT 263 278
 CC REPEAT 282 298
 CC REPEAT 306 317
 CC DISULFID 116 158
 CC DISULFID 147 194
 CC DISULFID 151 196
 CC DISULFID 141 141
 CC DISULFID 150 150
 CC CARBOHYD 160 160
 CC CARBOHYD 190 190
 CC CARBOHYD 292 292
 CC SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;
 Query Match 87.0%; Score 40; DB 1; Length 326;


```

DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A (VEGF-A) (Vascular permeability
DE factor) (VPF).
GN VEGF OR VEGFA.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognath; Caviidae; Cavia.
NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bile duct;
RA Berse B.;
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
CC cell growth. Induces endothelial proliferation and vascular
CC permeability (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84230; AAA37057.1; -
DR HSSP; P15692; 1VGH.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein.
FT DISULFID 25 67 BY SIMILARITY.
FT DISULFID 56 101 BY SIMILARITY.
FT DISULFID 60 103 BY SIMILARITY.
FT DISULFID 50 50 INTERCHAIN (BY SIMILARITY).
FT DISULFID 59 59 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 164 AA; 19330 MW; 9EB86A81A9D5DC44 CRC64;

Query Match 78.3%; Score 36; DB 1; Length 164;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
DB 60 CNDESLIC 67

RESULT 5
VEGA_BOVIN STANDARD; PRT; 190 AA.
AC P15691;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovine; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
RX MEDLINE=90069608; PubMed=2479985;
RA Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic
RT mitogen.";
RL Science 246:1306-1309(1989).
RN [2]
RP SEQUENCE OF 27-190 FROM N.A. (ISOFORMS ALPHA AND BETA).
RX MEDLINE=90121225; PubMed=2610687;
RA Tischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J.,
RA Lau K., Crisp T., Fiddes J.C., Abraham J.A.;
RT "Vascular endothelial growth factor: a new member of the platelet-
RT derived growth factor gene family.";
RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
RN [3]
RP SEQUENCE OF 27-31.
RX MEDLINE=89286596; PubMed=2735925;
RA Ferrara N., Henzel W.J.;
RT "Pituitary follicular cells secrete a novel heparin-binding growth
RT factor specific for vascular endothelial cells.";
RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: alpha (shown here) and beta; are
CC produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32976; AAA50502.1; -
DR EMBL; M31836; AAA50804.1; -
DR EMBL; M33750; AAA30805.1; -
DR PIR; A33255; A33255.
DR PIR; A33787; A33787.
DR PIR; B40080; B40080.
DR HSSP; P15692; 1VGH.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Alternative splicing; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 139 183 MISSING (IN ISOFORM BETA).
FT VARSPLIC 184 184 R -> K (IN ISOFORM BETA).
SQ SEQUENCE 190 AA; 22310 MW; EDBF903E46E24789 CRC64;

```

Query Match 78.3%; Score 36; DB 1; Length 190;
 Best Local Similarity 75.0%; Pred. No. 3.5;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESESLIC 8
 DB 86 CNESESLIC 93

RESULT 7
 VEGD_MOUSE STANDARD; PRT: 358 AA.
 AC P97946;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FtGF).
 GN FtGF OR VEGF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

SEQUENCE FROM N.A.
 STRAIN=C57BL/6J; TISSUE=Fibroblast;
 RC MEDLINE=97030254; PubMed=8876195;
 RA Orlandini M., Marconcin L., Ferruzzi R., Oliviero S.;
 RT "Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Lung;
 RC MEDLINE=97349118; PubMed=9205122;
 RA Yamaoka Y., Nezu J.-I., Shimane M., Hirata Y.;
 RT "Molecular cloning of a novel vascular endothelial growth factor, VEGF-D";
 RT Genomics 42:483-488(1997).
 RL [3]
 RN DEVELOPMENTAL STAGE.
 RP MEDLINE=98288130; PubMed=9622638;
 RA Avantsagiotou V., Orlandini M., Acampora D., Oliviero S., Simeone A.;
 RT "Embryonic expression pattern of the murine ftgf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family";
 RT Mech. Dev. 73:221-224(1998).
 RL [4]
 RN RECEPTOR SPECIFICITY.
 RP MEDLINE=21276411; PubMed=11279005;
 RA Baldwin M.E., Catimel B., Nice E.C., Roufall S., Hall N.E.,
 RA Stenvers K.L., Karkkainen M.J., Allitalo K., Stacker S.A., Achen M.G.;
 RT "The specificity of receptor binding by vascular endothelial growth factor-d is different in mouse and man";
 RT J. Biol. Chem. 276:19166-19171(2001).
 RL [5]
 RN FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor.
 CC SUBUNIT: Homodimer; non-covalent and antiparallel.
 CC TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.
 CC DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several body structures and organs of the embryo such as limb buds, acoustic ganglion, teeth, heart, anterior pituitary as well as lung and kidney mesenchyme, liver, derma, and perosteum of the vertebral column.

CC -1- INDUCTION: By the transcription factor c-fos.
 CC -1- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS) bound by non-covalent interactions (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch>).
 CC EMBL: X99572; CAA67892.1; -;
 CC EMBL: D89628; BAA14002.1; -;
 CC HSSP: P15692; IVP.
 CC MGD: MGI:108037; FtGF.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF_1.
 CC ProDom: PD001629; PDGF_1.
 CC SMART: SM00141; PDGF_1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS0278; PDGF_2; 1.
 CC DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
 CC KW Cleavage on pair of basic residues; Multigene family.
 CC SIGNAL 21
 CC PROPEP 22 93
 CC CHAIN 94 210
 CC PROPEP 211 358
 CC DOMAIN 227 323
 CC FT REPEAT 227 242
 CC FT REPEAT 263 278
 CC FT REPEAT 282 298
 CC FT REPEAT 306 323
 CC FT REPEAT 323 323
 CC FT DISULFID 116 158
 CC FT DISULFID 147 194
 CC FT DISULFID 151 196
 CC FT DISULFID 141 141
 CC FT DISULFID 150 150
 CC FT DISULFID 160 160
 CC FT CARBOHYD 190 190
 CC FT CARBOHYD 292 292
 CC SQ SEQUENCE 358 AA; 40908 MW; 6636B17BF07037C CRC64;

Query Match 78.3%; Score 36; DB 1; Length 358;
 Best Local Similarity 62.5%; Pred. No. 6.4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESESLIC 8
 DB 151 CNESESLIC 158

RESULT 8
 FASC_STRPU STANDARD; PRT: 495 AA.
 ID FASC_STRPU
 AC O05634;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fascin.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 CC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
 CC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=94022326; PubMed=8415664;
 RX Bryan J., Edwards R.H., Matsudaira P., Otto J., Wulfschlegel J.,
 RT "Pascin, an echinoid actin-binding protein, is a homolog of the
 RL Drosophila singed gene product.",
 RN Proc. Natl. Acad. Sci. U.S.A. 90:9115-9119(1993).
 RP CHARACTERIZATION.
 RX MEDLINE=79091184; PubMed=731692;
 RA Bryan J., Kane R.E.;
 RT "Separation and interaction of the major components of sea urchin
 RL actin gel.",
 RN J. Mol. Biol. 125:207-224(1978).
 RP CHARACTERIZATION.
 RX MEDLINE=77051438; PubMed=1033188;
 RA Kane R.E.;
 RT "Actin polymerization and interaction with other proteins in
 RL temperature-induced gelation of sea urchin egg extracts.",
 CC J. Cell Biol. 71:704-714(1976).
 CC -1- FUNCTION: ACTS AS AN ACTIN BUNDLING PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE FASCIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: L12047; AAC37183.1; -
 DR Actin-binding.
 KW FT INIT.MET 0
 SQ SEQUENCE 495 AA; 54813 MW; D3B0B5C52C21593 CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 1; Length 495;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
 Db 108 CNEESLIC 115

RESULT 9
 PKSL_BACSU STANDARD; PRT; 4427 AA.
 AC 003470;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative polyketide synthase (PKS).
 GN PKSL OR PKSX OR PKSA OR OUTG.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OK NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / PBI424;
 RX MEDLINE=93345824; PubMed=8344529;
 RA Gallizzi A., Albertini A.M., Perani P., Tognoni A., Grandi G.,
 RT "A Bacillus subtilis large ORF coding for a polypeptide highly
 RL similar to polyketide synthases.",
 RN Gene 130:65-71(1993).
 RP SEQUENCE OF 3619-4427 FROM N.A.
 RC STRAIN=168 / PBI424;
 RA Grandi G.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
 CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
 CC SECONDARY METABOLISM.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: 214098; CAA78479.1; -
 DR EMBL: U11039; AA85145.1; -
 DR EMBL: 235133; CAA84504.1; -
 DR EMBL: 299113; CABI3602.1; -
 DR PIR: S25021; S25021.
 DR HSSP: P27796; 1PXT.
 DR Subtilist; BG10698; PKSL.
 DR InterPro: IPR000794; ketoacyl_synth.
 DR InterPro: IPR003880; phosphopant_attach.
 DR Pfam: PF00109; ketoacyl_synth_4.
 DR Pfam: PF02801; ketoacyl_synth_4.
 DR Pfam: PF00550; pp-binding; 5.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 5.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE: PS50075; ACP DOMAIN; 5.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
 FT DOMAIN 211 280
 FT DOMAIN 382 759
 FT DOMAIN 937 1115
 FT DOMAIN 1409 1602
 FT DOMAIN 1687 1759
 FT DOMAIN 1876 2253
 FT DOMAIN 2491 2560
 FT DOMAIN 2632 2701
 FT DOMAIN 2823 3182
 FT DOMAIN 3575 3776
 FT DOMAIN 3854 3923
 FT DOMAIN 4019 4373
 FT BINDING 243 243
 FT BINDING 1723 1723
 FT BINDING 2523 2523
 FT BINDING 2664 2664
 FT BINDING 3886 3886
 SQ SEQUENCE 4427 AA; 493398 MW; 9612521E561AB9F2 CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 1; Length 4427;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
 Db 1205 CNEESLIC 1212

RESULT 10
 CPJL_RABIT STANDARD; PRT; 500 AA.
 AC P52766;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Cytochrome P450 201 (EC 1.14.14.1) (CYP201).
 GN CYP201.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

CC NCB1_TaxID=9986;
 CC (1)
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC MEDLINE-92011499; PubMed-1717443;
 CC Kikita Y., Sogawa K., Hanlu M., Kinoshita M., Kusunose E., Nojima Y.,
 CC Yamamoto S., Ichihara K., Kusunose M., Fujii-Kuriyama Y.,
 CC "A novel species of cytochrome P-450 (P-450ib) specific for the small
 CC intestine of rabbits cDNA cloning and its expression in COS cells."
 CC J. Biol. Chem. 266:17821-17825(1991).
 CC -1- FUNCTION: CATALYZES THE N-DEMEETHYLATION OF BENZOPRETHAMINE TO
 CC FORMALDEHYDE.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: SMALL INTESTINE.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90405; BAA14401.1; -
 CC HSSP: P00179; 1DT6.
 CC InterPro: IPR001128; Cyt_P450.
 CC Pfam: PF00067; P450.1.
 CC PRINTS: PR00385; P450.
 CC PROSITE: PS00086; CYTOCHROME_P450.1.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 CC Microsome; Endoplasmic reticulum.
 CC INT MET 0
 CC BINDING 446 446 HEME (BY SIMILARITY).
 CC SEQUENCE 500 AA; 57194 MW; A0EE628E5FF23DE CRC64;
 CC
 CC Query Match 73.9%; Score 34; DB 1; Length 500;
 CC Best Local Similarity 85.7%; Pred. No. 22;
 CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC OY 2 NEESLIC 8
 CC 111:111
 CC DB 296 NEENLIC 302
 CC
 CC RESULT 11
 CC CPJ6_MOUSE STANDARD; PRT; 501 AA.
 CC ID CPJ6_MOUSE
 CC AC 054750;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Cytochrome P450 2J6 (EC 1.14.14.1) (CYP1J6) (Arachidonic acid
 CC epoxygenase).
 CC GN CYP2J6.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 CC MEDLINE-98234557; PubMed-9570962;
 CC RA Ma J., Ramachandran S., Fiedorek F.T. Jr., Zeldin D.C.;
 CC "Mapping of the CYP2J cytochrome P450 genes to human chromosome 1 and
 CC mouse chromosome 4."
 CC RT Genomics 49:152-155(1998).
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U62295; AAB87636.1; -
 CC HSSP: P00179; 1DT6.
 CC MGD: MGI:1270148; CYP2J6.
 CC InterPro: IPR001128; Cyt_P450.
 CC Pfam: PF00067; P450.1.
 CC PRINTS: PR00385; P450.
 CC PROSITE: PS00086; CYTOCHROME_P450.1.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 CC Microsome; Endoplasmic reticulum.
 CC BINDING 447 447 HEME (BY SIMILARITY).
 CC SEQUENCE 501 AA; 57820 MW; 4E5A26B35D831E37 CRC64;
 CC
 CC Query Match 73.9%; Score 34; DB 1; Length 501;
 CC Best Local Similarity 85.7%; Pred. No. 22;
 CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC OY 2 NEESLIC 8
 CC 111:111
 CC DB 296 NEENLIC 302
 CC
 CC RESULT 12
 CC CPJ3_RAT STANDARD; PRT; 502 AA.
 CC ID CPJ3_RAT
 CC AC P51590;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE Cytochrome P450 2J3 (EC 1.14.14.1) (CYP1J3).
 CC GN CYP2J3.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=FISCHER 344;
 CC MEDLINE-97284730; PubMed-9139707;
 CC Wu S., Chen W., Murphy E., Gabel S., Tomer K.B., Foley J.,
 CC Steenbergen C., Falck J.R., Moormaw C.R., Zeldin D.C.;
 CC "Molecular cloning, expression, and functional significance of a
 CC cytochrome P450 highly expressed in rat heart myocytes."
 CC J. Biol. Chem. 272:12551-12559(1997).
 CC -1- FUNCTION: THIS ENZYME METABOLIZES ARACHIDONIC ACID PREDOMINANTLY
 CC VIA A NADPH-DEPENDENT OLEFIN EPOXYDATION MAINLY TO 14,15-, 11,12-,
 CC AND 8,9-EPOXYEICOSATRIENOIC ACIDS (EET). IT ALSO ACTS AS AN OMEGA-
 CC 1-HYDROXYLASE BY METABOLIZING ARACHIDONIC ACID TO 19-
 CC HYDROXYEICOSATETRAENOIC ACID (19-OH-AA).
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN HEART AND LIVER.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U39943; AAB48545.1; -

DR HSSP: P00179; 1DT6.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
DR OXIDOREDUCTASE: Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; NADP.
FT BINDING 448 448 HEME (BY SIMILARITY).
SQ SEQUENCE 502 AA; 57969 MW; CCCC0AE67977FEF31 CRC64;

Query Match
Best Local Similarity 85.7%; Score 34; DB 1; Length 502;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 NEESLIC 8
Db 297 NEESLIC 303

RESULT 13
RBB2_HUMAN STANDARD; PRT: 1722 AA.
ID RBB2_HUMAN
AC P29375;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinoblastoma-binding protein 2 (RBBP-2).
GN RBBP2 OR RBP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:94020841; PubMed-8414517;
RA Fattley A.R., Helin K., Dembski M.S., Dyson N., Harlow E.,
RA Vuocolo G.A., Hanobik M.G., Haskell K.M., Oliff A., Defeo-Jones D.,
RT Jones R.E.;
RT "Characterization of the retinoblastoma binding proteins RBP1 and
RBP2.";
RL Oncogene 8:3149-3156(1993).
RN [2]
RP SEQUENCE OF 1102-1569 FROM N.A.
RX MEDLINE-91312450; PubMed-1857421;
RA Defeo-Jones D., Huang P.S., Jones R.E., Haskell K.M., Vuocolo G.A.,
RA Hanobik M.G., Huber H.E., Oliff A.;
RT "Cloning of cDNAs for cellular proteins that bind to the
retinoblastoma gene product.";
RL Nature 352:251-254(1991).
CC -1- FUNCTION: INTERACTS WITH THE VIRAL PROTEIN-BINDING DOMAIN OF THE
RETINOBLASTOMA PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: TO THE C-TERMINAL DOMAIN OF THE XE169 PROTEIN AND
THE JUMONJI PROTEIN.
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: S66431; AAR28544.1; -.
DR PTR: S16954; S16954.
DR MIM: 180202; -.
DR InterPro: IPR001606; ARID.
DR InterPro: IPR003347; JmjN.
DR InterPro: IPR003349; JmjN.
DR InterPro: IPR001965; PHD.
DR InterPro: IPR004198; Zf-C5HC2.
DR Pfam: PF01388; ARID; 1.

DR Pfam: PF02373; JmjC; 1.
DR Pfam: PF02375; JmjN; 1.
DR Pfam: PF00628; PHD; 3.
DR Pfam: PF02928; Zf-C5HC2; 1.
DR SMART: SM00501; BRIGH7; 1.
DR SMART: SM00545; JmjN; 1.
DR SMART: SM00249; PHD; 3.
KW Trans-acting factor; Nuclear protein.
FT CONFLICT 1563 1566 MISSING (IN REF. 2).
SQ SEQUENCE 1722 AA; 195815 MW; BCFB8A8BAE69A652 CRC64;

Query Match
Best Local Similarity 62.5%; Score 34; DB 1; Length 1722;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NEESLIC 8
Db 692 CNPRLVC 699

RESULT 14
VEGA_MOUSE STANDARD; PRT: 214 AA.
ID VEGA_MOUSE
AC 000731;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM VEGF-1; VEGF-2 AND VEGF-3).
RX MEDLINE-92274860; PubMed-1592003;
RA Breier G., Albrecht U., Sterrer S., Risau W.;
RT "Expression of vascular endothelial growth factor during embryonic
angiogenesis and endothelial cell differentiation.";
RL Development 114:521-532(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM VEGF-1).
RX MEDLINE-92355593; PubMed-1644816;
RA Claffey K.P., Wilkison W.O., Spiegelman B.M.;
RT "Vascular endothelial growth factor. Regulation by cell
differentiation and activated second messenger pathways.";
RL J. Biol. Chem. 267:16317-16322(1992).
RN [3]
RP SEQUENCE OF 1-3 FROM N.A.
RX MEDLINE-96216498; PubMed-8632007;
RA Shima D.T., Kuroki M., Deutsch U., Ng Y., Adams A.P., D'Amore P.A.;
RT "The mouse gene for vascular endothelial growth factor. Genomic
structure, definition of the transcriptional unit, and
characterization of transcriptional and post-transcriptional
regulatory sequences.";
RL J. Biol. Chem. 271:3877-3883(1996).
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
endothelial cell growth. It induces endothelial cell
proliferation, promotes cell migration, inhibits apoptosis, and
induces permeabilization of blood vessels. It binds to the
VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: VEGF-1 and VEGF-2 are secreted while VEGF-3
remains cell-surface associated unless released by heparin.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms: VEGF-3/VEGF188 (shown here),
VEGF-1/VEGF164 and VEGF-2/VEGF120; are produced by alternative
splicing.
CC -1- TISSUE SPECIFICITY: In developing embryos, expressed mainly in the
choroid plexus, paraventricular neuroepithelium, placenta and

CC kidney glomeruli. Also found in bronchial epithelium, adrenal
 CC gland and in seminiferous tubules of testis. High expression of
 CC VEGF continues in kidney glomeruli and choroid plexus in adults.
 CC -i- DOMAIN: VEGF-3 contains a basic insert which acts as a cell
 CC retention signal.
 CC -i- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S37052; AAB22252.1; -;
 CC EMBL: S38083; AAB22253.1; -;
 CC EMBL: S38100; AAB22254.1; -;
 CC EMBL: M95200; AAA0547.1; -;
 CC EMBL: U41383; -; NOT_ANNOTATED_CDS.
 CC PIR: A43351; A43351.
 CC HSSP: P15692; 2VPF.
 CC MGD: MGI:103178; Vegf.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF; 1.
 CC ProDom: PD001629; PDGF; 1.
 CC SMART: SM00141; PDGF; 1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC DR PROSITE: PS0278; PDGF_2; 1.
 CC KM Mitogen: Angiogenesis; Growth factor; Glycoprotein; Signal;
 CC Hepatin-binding; Alternative splicing; Multigene family.
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 214
 CC FT DISULFID 51 93
 CC FT DISULFID 82 127
 CC FT DISULFID 86 129
 CC FT DISULFID 76 76
 CC FT DISULFID 85 85
 CC FT CARBOHYD 100 100
 CC FT VARSPLIC 140 140
 CC FT VARSPLIC 141 164
 CC FT VARSPLIC 141 208
 CC FT CONFLICT 117 118
 CC SQ SEQUENCE 214 AA; 25283 MW; B5540B51E4BB6E17 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 214;
 Best Local Similarity 62.58; Pred. NO. 15;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 ||:|:|
 DB 86 CNDEALEC 93

RESULT 15
 VEGF_RAT STANDARD: PRT: 214 AA.
 AC P16612; O90XG7; O90XG6; O91KX7;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
 DE permeability factor) (VPF).
 GN VEGF OR VEGFA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM VEGF-A164), AND SEQUENCE OF 27-190.
 RA MEDLINE=90207249; PubMed=2320579; Kwok P.W., Sullivan K.A.,
 RA Conn G., Bayne M.L., Soderman D.D.,

RA Palisi T.M., Hope D.A., Thomas K.A.;
 RT "Amino acid and cDNA sequences of a vascular endothelial cell mitogen
 RT that is homologous to platelet-derived growth factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2628-2633(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS VEGF-A188; VEGF-A164; VEGF-A144 AND
 RP VEGF-A120).
 RA Ishii H., Arakawa T., Okayama M., Oota I., Takuma T., Inomata K.;
 RT "Developmental expression of vascular endothelial growth factor-A
 RT (VEGF-A) splicing variants, VEGF-A188, VEGF-A164, and VEGF-A120 in rat
 RT masseter muscle.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 27-40.
 RC TISSUE-Glial tumor;
 RX MEDLINE=95221439; PubMed=7706320;
 RA Soderman D.D., Palisi T.M., Sullivan K.A., Thomas K.A.;
 RT "Purification and characterization of a naturally occurring vascular
 RT endothelial growth factor, placenta growth factor heterodimer.";
 RL J. Biol. Chem. 270:7717-7723(1995).
 CC -i- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
 CC endothelial cell growth. It induces endothelial cell
 CC proliferation, promotes cell migration, inhibits apoptosis, and
 CC induces permeabilization of blood vessels. It binds to the
 CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
 CC heparin (By similarity).
 CC -i- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
 CC with PlGF (By similarity).
 CC -i- SUBCELLULAR LOCATION: VEGF-A120 is acidic and freely secreted.
 CC VEGF-A164 is more basic, has heparin-binding properties and,
 CC although a significant proportion remains cell-associated, most is
 CC freely secreted. VEGF-A188 is very basic; it is cell-associated
 CC after secretion and is bound avidly by heparin and the
 CC extracellular matrix, although it may be released as a soluble
 CC form by heparin, heparinase or plasmin (By similarity).
 CC -i- ALTERNATIVE PRODUCTS: At least 4 isoforms: VEGF-A188 (shown here),
 CC VEGF-A164, VEGF-A144 and VEGF-A120; are produced by alternative
 CC splicing.
 CC -i- TISSUE SPECIFICITY: Expressed in the pituitary, in brain, in
 CC particularly in suprapubic and paraventricular nuclei and the
 CC choroid plexus. Also found abundantly in the corpus luteum of
 CC the ovary and in kidney glomeruli.
 CC -i- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M32167; AAA41211.1; -;
 CC EMBL: AF215725; AAF19211.1; -;
 CC EMBL: AF215726; AAF19212.1; -;
 CC EMBL: AF222779; AAF25958.1; -;
 CC PIR: A35987; A35987.
 CC HSSP: P15692; 1VPF.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF; 1.
 CC ProDom: PD001629; PDGF; 1.
 CC SMART: SM00141; PDGF; 1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC DR PROSITE: PS0278; PDGF_2; 1.
 CC KM Mitogen: Angiogenesis; Growth factor; Glycoprotein; Signal;
 CC Hepatin-binding; Alternative splicing; Multigene family.
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 214
 CC FT DISULFID 51 93
 CC FT DISULFID 82 127
 CC FT DISULFID 86 129
 CC FT DISULFID 76 76
 CC VEGF_RAT STANDARD: PRT: 214 AA.
 AC P16612; O90XG7; O90XG6; O91KX7;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
 DE permeability factor) (VPF).
 GN VEGF OR VEGFA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM VEGF-A164), AND SEQUENCE OF 27-190.
 RA MEDLINE=90207249; PubMed=2320579; Kwok P.W., Sullivan K.A.,
 RA Conn G., Bayne M.L., Soderman D.D.,

FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .).
 FT VARSPLIC 140 140 K -> N (IN ISOFORM VEGF-A164).
 FT VARSPLIC 141 164 MISSING (IN ISOFORM VEGF-A164).
 FT VARSPLIC 141 208 MISSING (IN ISOFORM VEGF-A120).
 FT VARSPLIC 165 208 MISSING (IN ISOFORM VEGF-A144).
 FT CONFLICT 101 101 V -> A (IN REF. 2; AAF19212).
 SQ SEQUENCE 214 AA; 25239 MW; 60FB876F5304946 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 214;
 Best Local Similarity 62.5%; Pred. No. 15;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CNESLIC 8
 11:111
 Db 86 CNDEALEC 93

Search completed: June 17, 2002, 16:24:26
 Job time: 1555 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:55:11 ; Search time 95.43 Seconds
(without alignments)
13.090 Million cell updates/sec

Title: US-09-761-636A-5

Perfect score: 72

Sequence: 1 CASELGKSTNFC 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	59.7	104	A39035	ribonuclease-relat
2	41	56.9	397	AC1574	acetate kinase hom
3	41	56.9	397	AH1220	hypothetical prote
4	40	55.6	383	T17722	Kruppel-like prote
5	39	54.2	409	S42384	hypothetical prote
6	39	54.2	1117	C85018	membrane-associate
7	38	52.8	378	G97177	NAC2-like protein
8	38	52.8	469	T46230	lycB protein (limp
9	37	51.4	319	A84947	XPMK2 protein - At
10	37	51.4	421	S53818	pectinesterase (EC
11	37	51.4	433	S70914	protein kinase C (
12	37	51.4	988	S35362	probable membrane
13	37	51.4	1174	T43051	hypothetical prote
14	37	51.4	1294	S77690	T-cell receptor be
15	36.5	50.7	230	S09778	hypothetical prote
16	36	50.0	35	B49046	probable alpha-ox
17	36	50.0	282	A84341	hypothetical prote
18	36	50.0	372	E85170	hypothetical prote
19	36	50.0	419	S34421	GTP-binding regula
20	36	50.0	442	H71410	probable RNA helic
21	36	50.0	483	S37055	catalase (EC 1.11.
22	36	50.0	683	S44780	C30A5.5 protein -
23	36	50.0	814	A71419	probable alpha-ox
24	36	50.0	944	T41711	hypothetical prote
25	36	50.0	1226	S69078	MG243 homolog H91
26	35.5	49.3	224	S73823	germin homolog F21
27	35.5	49.3	266	T01199	viral capsid prote
28	35	48.6	47	JC6158	ORF MSV185 hypothe
29	35	48.6	62	T28346	

ALIGNMENTS

30	35	48.6	66	2	A27406	toxin CstII - Mexi
31	35	48.6	80	2	PC4272	cdus protein - Clo
32	35	48.6	155	2	DB2152	conserved hypothet
33	35	48.6	182	2	H71899	hypothetical prote
34	35	48.6	207	2	AE0199	probable lipoprote
35	35	48.6	218	2	T33545	hypothetical prote
36	35	48.6	228	2	S67199	hypothetical prote
37	35	48.6	238	2	S77699	inner cell wall ma
38	35	48.6	265	2	AF2007	hypothetical prote
39	35	48.6	276	2	S27641	hypothetical prote
40	35	48.6	308	2	AF1990	hypothetical prote
41	35	48.6	375	2	H82329	alanine racemase,
42	35	48.6	393	2	E64239	acetate kinase (EC
43	35	48.6	396	2	D81339	acetate kinase (EC
44	35	48.6	415	2	A11611	competence-damage
45	35	48.6	483	1	T43234	protein kinase (EC

RESULT 1

A39035 ribonuclease-related anti-tumor protein - northern leopard frog (fragment)

C:Species: Rana pipiens (northern leopard frog)

C:Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 30-Jun-1993

C:Accession: A39035

R:Rardelt, W.; Mikulski, S.M.; Shogen, K.

J. Biol. Chem. 266, 245-251, 1991

A:Title: Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and

A:Reference number: A39035; M0ID:91093131

A:Accession: A39035

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-104 <ARD>

C:Superfamily: pancreatic ribonuclease

Query Match 59.7%; Score 43; DB 2; Length 104;

Best local similarity 61.5%; Pred. No. 1.8;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNFC 13

DB 75 CKYKLRKSTNFC 87

RESULT 2

AC1574 acetate kinase homolog AckA2 [imported] - Listeria innocua (strain Clp11262)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AC1574

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; B

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsi

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A

ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; We

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; M0ID:21537279; PMID:11679669

A:Accession: AC1574

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96363.1; PID:G16413591; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: AckA2

C:Superfamily: acetate kinase

Query Match 56.9%; Score 41; DB 2; Length 397;

Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CASELGKSTNTP 11
|||
Db 214 CAIEAGKSVNT 224

RESULT 3

AH1220

acetate kinase homolog AckA2 [imported] - *Listeria monocytogenes* (strain EGD-e)
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C/Accession: AH1220

R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H. Science 294, 849-852, 2001.

A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Ma Ok, C.; Schlueter, T.; Simoes, N.; Tlertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AH1220

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-397 <GUA>

A/Cross-references: GB:NC_003210; PIDN:CAC99246.1; PID:g16410584; GSPDB:GN00177

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: AckA2

C/Superfamily: acetate kinase

Query Match 56.9%; Score 41; DB 2; Length 397;
Best Local Similarity 72.7%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CASELGKSTNTP 11
|||
Db 214 CAIEAGKSVNT 224

RESULT 4

T17722

hypothetical protein A231L - *Chlorella* virus PCV-1
C/Species: *Chlorella* virus PCV-1

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T17722

R/Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999

A/Reference number: Z18806
A/Accession: T17722

A/Status: preliminary

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-383 <GRA>

A/Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC9559.1
A/Experimental source: specific host *Chlorella* strain NC64A

C/Genetics:

A/Note: A231L

A/Note: A231L

Query Match 55.6%; Score 40; DB 2; Length 383;
Best Local Similarity 46.2%; Pred. No. 21;

Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 CASELGKSTNTPC 13
|||
Db 304 CVMNIGKKKNEFC 316

RESULT 5

S42384

Kruppel-like protein - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-Sep-1997

C/Accession: S42384

R/Smith, A.

submitted to the EMBL Data Library, March 1994
A/Reference number: S42368

A/Accession: S42384

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-409 <SMI>
A/Cross-references: EMBL:230662; NID:g459581; PID:g459587

C/Genetics:

A/Introns: 10/3; 48/3; 182/3; 315/3; 360/3

Query Match 54.2%; Score 39; DB 2; Length 409;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGKSTNTP 12
|||
Db 229 CKSDLSPQTQF 240

RESULT 6

C85018

hypothetical protein AT4901400 [imported] - *Arabidopsis thaliana*

C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C/Accession: C85018

R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold
Nature 402, 769-777, 1999

A/Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.
A/Reference number: AB5001; MUID:20083468

A/Accession: C85018

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1117 <STO>
A/Cross-references: GB:NC_001268; NID:g7267637; PIDN:CAB80949.1; GSPDB:GN00140

C/Genetics:

A/Gene: AT4901400

A/Map position: 4

Query Match 54.2%; Score 39; DB 2; Length 1117;
Best Local Similarity 58.3%; Pred. No. 87;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CASELGKSTNTP 12
|||
Db 922 CSELGELSTF 933

RESULT 7

G97177

membrane-associated sensory histidine kinase-like Arpase [imported] - *Clostridium acetobutylicum*

C/Species: *Clostridium acetobutylicum*
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C/Accession: G97177

R/Nolling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;
Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: G97177

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-378 <KUR>

A/Cross-references: GB:AE001437; PIDN:AAK80210.1; PID:g15025255; GSPDB:GN00168
A/Experimental source: *Clostridium acetobutylicum* ATCC824

C/Genetics:

A/Gene: CAC2253

Query Match
Best Local Similarity 52.8%; Score 38; DB 2; Length 378;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CASELGKSTNTF 11
|||: |||: |
DB 278 CASDIKSTTT 288

RESULT 8

T46230
MAC2-like protein - Arabidopsis thaliana
N:Alternate names: protein T9C5.120
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46230
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the Protein Sequence Database, December 1999
A:Reference number: 223026
A:Accession: T46230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <RTE>
A:Cross-references: EMBL:AL132964
A:Experimental source: cultivar Columbia; BAC clone T9C5
C:Genetics:
A:Map position: 3
A:Introns: 60/1; 153/3; 233/3; 332/3; 402/3
A:Note: T9C5.120

Query Match
Best Local Similarity 52.8%; Score 38; DB 2; Length 469;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 SELGKSTNTF 12
||||: |||: |
DB 260 SELGKSTNTF 269

RESULT 9

AB4947
lyb protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: AB4947
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A:Reference number: AB4930; MUID:20445173
A:Accession: AB4947
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <STO>
A:Cross-references: GB:AP000398; GSPDB:GM00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: lyb; BU147
C:Superfamily: penicillin tolerance protein

Query Match
Best Local Similarity 51.4%; Score 37; DB 2; Length 319;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 SELGKSTNTF 12
||||: |||: |
DB 232 AELGKETGTF 241

RESULT 10
S53818

xpmc2 protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jul-2000
C:Accession: S53818
R:Su, J.Y.; Maller, J.L.
Mol. Gen. Genet. 246, 387-396, 1995
A:Title: Cloning and expression of a Xenopus gene that prevents mitotic catastrophe
A:Reference number: S53818; MUID:95157530
A:Accession: S53818
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-421 <SUJ>
A:Cross-references: GB:U10185; NID:9595379; PIDN:AAA82179.1; PID:9595380

Query Match
Best Local Similarity 51.4%; Score 37; DB 2; Length 421;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 CASELGKSTNTF 11
||: |||: |
DB 13 CAGSLGKTANT 23

RESULT 11

S70914
pectinesterase (EC 3.1.1.11) precursor - Erwinia chrysanthemi
N:Alternate names: pectin methyltransferase
C:Species: Erwinia chrysanthemi
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S70914; S70913
R:Shevchik, V.E.; Condemine, G.; Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J.
submitted to the EMBL Data Library, February 1995
A:Description: Characterization of pectin methyltransferase B, an outer membrane lipoprotein
A:Reference number: S70914
A:Accession: S70914
A:Molecule type: DNA
A:Residues: 1-433 <SHE>
A:Cross-references: EMBL:X84665; NID:91212890; PIDN:CAA59151.1; PID:91212891
R:Shevchik, V.E.; Condemine, G.; Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J.
Mol. Microbiol. 19, 455-466, 1996
A:Title: Characterization of pectin methyltransferase B, an outer membrane lipoprotein
A:Reference number: S70913; MUID:96228695
A:Accession: S70913
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-53; 234-361 <SHA>
A:Cross-references: EMBL:X84665; NID:91212890; PIDN:CAA59151.1
A:Experimental source: strain 3937
C:Genetics:
A:Gene: pectB
A:Start codon: GTG
C:Superfamily: pectinesterase pectB
C:Keywords: carboxylic ester hydrolase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-43/Product: pectinesterase #status predicted <MAT>

Query Match
Best Local Similarity 51.4%; Score 37; DB 2; Length 433;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFC 13
||: | : |||: |
DB 192 CATKAGATINTTC 204

RESULT 12

S33362
protein kinase C (EC 2.7.1.1-) pck1 - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: pombe C-kinase 1
C:Species: Schizosaccharomyces pombe
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 06-Oct-2000

C:Accession: S35362; T37866; T38203
 R:Toda, T.; Shimanuki, M.; Yanagida, M.
 EMBL J. 12, 1987-1995, 1993
 A:Title: Two novel protein kinase C-related genes of fission yeast are essential for cell growth
 A:Reference number: S35362; MUID:93259141
 A:Accession: S35362
 A:Molecule type: DNA
 A:Residues: 1-988 <TOD>
 A:Cross-references: EMBL:D14337; NID:q303938; PIDN:BAA03267.1; PID:q303939
 R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsby, S.V.
 Submitted to the EMBL Data Library, February 1996
 A:Reference number: 221750
 A:Accession: T37866
 A:Molecule type: DNA
 A:Residues: 93-988 <DEV>
 A:Cross-references: EMBL:Z69795; PIDN:CAA93697.1; GSPDB:GN00066; SPDB:SPAC1768.14C
 R:Gishchuk, K.; McIntosh, J.R.; Devlin, K.; Churcher, C.; Barrell, B.G.; Rajandream, M.
 Submitted to the EMBL Data Library, February 1996
 A:Reference number: 221778
 A:Accession: T38203
 A:Molecule type: DNA
 A:Residues: 1-26, 'AMVASTKRP', 36-137 <GR1>
 A:Cross-references: EMBL:Z69730; PIDN:CAA93602.1; GSPDB:GN00066; SPDB:SPAC22H10.01C
 A:Experimental source: strain 972h-; cosmid c22H10
 C:Genetics:
 A:Gene: pck1
 A:Map position: 1L
 A:Introns: 767/2; 834/3; 897/3; 959/1
 C:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology; pck1
 C:Keywords: ATP; duplication; phospholipid binding; phosphotransferase; serine/threonine
 F:414-461/Domain: protein kinase C zinc-binding repeat homology <K21>
 F:481-530/Domain: protein kinase C zinc-binding repeat homology <K22>
 F:562-923/Domain: protein kinase homology <K22>
 F:670-678/Region: protein kinase ATP-binding motif

Query Match 51.4%; Score 37; DB 1; Length 988;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 LGKSTNFC 13
 |||
 Db 817 LGNTSTFC 825

RESULT 13
 T43051
 protein kinase C (EC 2.7.1.-) PKC1 - fungus (Cochliobolus heterostrophus)
 C:Species: Cochliobolus heterostrophus, Bipolaris maydis
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43051
 R:Oester, B.M.
 FEMS Microbiol. Lett. 165, 273-280, 1998
 A:Title: PKC1, encoding a protein kinase C, and FAT1, encoding a fatty acid transporter
 A:Reference number: Z22303; MUID:98415124
 A:Accession: T43051
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1174 <OES>
 A:Cross-references: EMBL:Y15839; NID:q2687848; PIDN:CAA75801.1; PID:q2687849
 A:Experimental source: strain C2 (ATCC 48329)
 C:Genetics:
 A:Gene: pck1
 A:Introns: 146/1; 234/1; 559/1; 952/2; 1019/3; 1155/1
 C:Superfamily: yeast protein kinase C; protein kinase C zinc-binding repeat homology; pck1
 C:Keywords: ATP; phospholipid binding; phosphotransferase; serine/threonine-specific protein kinase
 F:459-506/Domain: protein kinase C zinc-binding repeat homology <K2N1>
 F:527-576/Domain: protein kinase C zinc-binding repeat homology <K2N2>

Query Match 51.4%; Score 37; DB 2; Length 1174;
 Best Local Similarity 46.7%; Pred. No. 2.1e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

OY 1 CASEL--GKSTNFC 13
 |||
 Db 996 CKEMWGTSTSTFC 1010

RESULT 14
 S77690
 probable membrane protein YOL075c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O1125; hypothetical protein O1130; hypothetical protein O1135
 C:Species: Saccharomyces cerevisiae
 C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 02-Mar-2001
 C:Accession: S77690; S66767; S66768
 R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66756
 A:Accession: S77690
 A:Molecule type: DNA
 A:Residues: 1-1294 <ALE>
 A:Cross-references: EMBL:Z74816; MIPS:YOL075c
 A:Note: this is a revision to the sequence from reference S66756
 A:Accession: S66767
 A:Molecule type: DNA
 A:Residues: 1-179, 'TTRGVFLVVKRED' <ALM>
 A:Cross-references: EMBL:Z74816
 A:Experimental source: strain S288C
 A:Note: this sequence has been revised in reference S77690
 A:Accession: S66768
 A:Molecule type: DNA
 A:Residues: 200-1294 <ALF>
 A:Cross-references: EMBL:Z74817
 A:Experimental source: strain S288C
 A:Note: this sequence has been revised in reference S77690
 A:Note: this sequence has been revised in reference S77690
 C:Genetics:
 A:Map position: 15L
 A:Note: YOL075c
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
 F:45-263/Domain: ATP-binding cassette homology <ABC1>
 F:62-69/Region: nucleotide-binding motif A (P-loop)
 F:376-392/Domain: transmembrane #status predicted <TM1>
 F:469-485/Domain: transmembrane #status predicted <TM2>
 F:496-512/Domain: transmembrane #status predicted <TM3>
 F:606-622/Domain: transmembrane #status predicted <TM4>
 F:710-916/Domain: ATP-binding cassette homology <ABC2>
 F:727-734/Region: nucleotide-binding motif A (P-loop)
 F:1042-1058/Domain: transmembrane #status predicted <TM5>
 F:1125-1141/Domain: transmembrane #status predicted <TM6>
 F:1177-1193/Domain: transmembrane #status predicted <TM7>
 F:1269-1285/Domain: transmembrane #status predicted <TM8>

Query Match 51.4%; Score 37; DB 2; Length 1294;
 Best Local Similarity 58.3%; Pred. No. 2.3e+02;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CASEL--GKSTNFC 12
 |||
 Db 1161 CGERLGMTNFC 1172

RESULT 15
 S09778
 hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)
 C:Species: human cytomegalovirus, human herpesvirus 5
 A:Note: host Homo sapiens (man)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
 C:Accession: S09778
 R:Chen, W.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Hornselli, M.; Barrell, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039
A:Accession: S09778
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-230 <CHE>
A:Cross-references: EMBL:X17403; NID:959591; PIDN:CA35448.1; PID:959620
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Superfamily: human cytomegalovirus hypothetical protein UL16
C:Keywords: glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-230/Product: hypothetical protein UL16 #status predicted <MAT>
F:189-206/Domain: transmembrane #status predicted <TM>
F:35,41,68,84,95,101,132,145/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.7%; Score 36.5; DB 2; Length 230;
Best Local Similarity 57.1%; Pred. No. 56;
Matches 8; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 CASEIG-KSTNTEFC 13
DB 25 CAVDLGSKSSNSTC 38

Search completed: June 17, 2002, 15:58:18
Job time: 187 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:31 ; Search time 44.84 Seconds
(without alignments)

11.226 Million cell updates/sec

Title: US-09-761-636A-5

Sequence: 1 CASP1GKSTNTEC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	75.0	354	1	VEGD_HUMAN
2	51	70.8	326	1	VEGD_RAT
3	51	70.8	358	1	VEGD_MOUSE
4	43	59.7	104	1	RN30_RANPI
5	39	54.2	409	1	YV5_CAEEL
6	37	51.4	319	1	LYTB_BUCAL
7	37	51.4	407	1	RPE2_HUMAN
8	37	51.4	433	1	PMB1_ERMCH
9	37	51.4	988	1	PK1_SCHPO
10	37	51.4	1174	1	KPCL_COCH
11	37	51.4	1294	1	YOH5_YEAST
12	37	50.7	230	1	UL16_HCMVA
13	36.5	50.0	220	1	WTH_HUMAN
14	36	50.0	282	1	BCA_HALNL
15	36	50.0	483	1	ESR2_SYRV
16	36	50.0	673	1	ESR2_MICUN
17	36	50.0	683	1	YK86_CAEEL
18	35.5	49.3	224	1	Y243_MYCPN
19	35.5	49.3	2156	1	RPL_HUMAN
20	35	48.6	66	1	SCX2_CENSU
21	35	48.6	238	1	CM14_YEAST
22	35	48.6	393	1	ACRA_MYCGE
23	35	48.6	602	1	TRAN_ECOLI
24	35	48.6	758	1	PARC_RHIME
25	35	48.6	855	1	ENV_HVIA2
26	35	48.6	980	1	YRS5_CAEEL
27	35	48.6	1041	1	CHS1_CRYNE
28	35	48.6	1151	1	KPCL_YEAST
29	35	48.6	1517	1	GLTB_ECOLI
30	35	48.6	1941	1	UBR1_KULIA
31	34.5	47.9	245	1	IF6_YEAST
32	34.5	47.9	300	1	SP18_YEAST
33	34.5	47.9	609	1	PEPE_BACSU

ALIGNMENTS

RESULT	ID	VEGD_HUMAN	STANDARD	PRT	354 AA.
AC	043915	VEGD_HUMAN	STANDARD	PRT	354 AA.
DT	01-MAR-2002	(rel. 41, Created)			
DT	01-MAR-2002	(rel. 41, Last sequence update)			
DT	01-MAR-2002	(rel. 41, Last annotation update)			
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).				
DE	FIGF OR VEGFD.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;				
RC	MEDLINE=97349118; PubMed=9205122;				
RX	Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;				
RA	"Molecular cloning of a novel vascular endothelial growth factor,				
RT	VEGF-D.";				
RL	Genomics 42:483-488(1997).				
RN	[2]	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;				
RC	MEDLINE=98140120; PubMed=9479493;				
RX	Rocchigliani M., Lestingi M., Luddi A., Orlandini M., Franco B.,				
RA	Rossi E., Ballabio A., Zuffardi O., Oliviero S.;				
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1				
RL	between the FIGF and the GRPR genes.";				
RN	Genomics 47:207-216(1998).				
RN	[3]	SEQUENCE FROM N.A.			
RP	MEDLINE=98118549; PubMed=9435229;				
RX	Achen M.G., Jeltsch M., Kuk E., Maekinen T., Vitalli A., Wilks A.F.,				
RA	Allitalo K., Stacker S.A.;				
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the				
RL	tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flk4).";				
RN	Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).				
RN	[4]	PROCESSING, AND SEQUENCE OF 89-94; 100-105 AND 206-213.			
RP	MEDLINE=20011413; PubMed=10542248;				
RX	Stacker S.A., Stenvers K.L., Caesar C., Vitalli A., Domagala T.,				
RA	Nice E.C., Rouffail S., Simpson R.J., Moritz R., Karpanen T.,				
RT	Allitalo K., Achen M.G.;				
RL	"Biosynthesis of vascular endothelial growth factor-D involves				
RT	proteolytic processing which generates non-covalent homodimers.";				
RN	J. Biol. Chem. 274:32127-32136(1999).				
RL	-I- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis				
CC	and endothelial cell growth, stimulating their proliferation and				
CC	migration and also has effects on the permeability of blood				
CC	vessels. May function in the formation of the venous and lymphatic				
CC	vascular systems during embryogenesis, and also in the maintenance				
CC	of differentiated lymphatic endothelium in adults. Binds and				
CC	activates VEGFR-2 (Flk1) and VEGFR-3 (Flk4) receptors.				
CC	-I- SUBUNIT: Homodimer; non-covalent and antiparallel.				

34	34.5	47.9	745	1	RNR_BUCAL	P57628 buchnera ap
35	34	47.2	122	1	UPPI_PIG	O29100 sus scrofa
36	34	47.2	335	1	CRL1_XENLA	O91574 xenopus lae
37	34	47.2	420	1	PEL_BACSU	P39116 bacillus su
38	34	47.2	430	1	KNT1_RAT	P01048 rattus norv
39	34	47.2	430	1	KNT2_RAT	P08932 rattus norv
40	34	47.2	465	1	COBO_ACICA	O43989 actinobact
41	34	47.2	474	1	2256_HUMAN	O9y2p7 homo sapien
42	34	47.2	533	1	YN4_YEAST	P48565 saccharomyc
43	34	47.2	823	1	SCH9_YEAST	P11792 saccharomyc
44	34	47.2	852	1	ENV_HVIA3	P19549 human immun
45	34	47.2	988	1	CICL1_HUMAN	P33523 homo sapien

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in lung, heart, small
CC intestine and fetal lung, and at lower levels in skeletal muscle,
CC colon, and pancreas.
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D96630; BAA24264.1; -
DR EMBL: Y12863; CAAT7370.1; -
DR EMBL: Y12864; CAAT7371.1; -
DR EMBL: Y12865; CAAT7371.1; JOINED.
DR EMBL: Y12866; CAAT7371.1; JOINED.
DR EMBL: Y12867; CAAT7371.1; JOINED.
DR EMBL: Y12868; CAAT7371.1; JOINED.
DR EMBL: Y12869; CAAT7371.1; JOINED.
DR EMBL: Y12870; CAAT7371.1; JOINED.
DR EMBL: A000185; CAA03942.1; -
DR MIM: 300091; -
DR HSSP: P15692; IVP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 88
FT CHAIN 89 205
FT PROPEP 206 354
FT DOMAIN 222 318
FT REPEAT 222 237
FT REPEAT 258 273
FT REPEAT 277 293
FT REPEAT 301 318
FT DISULFID 111 153
FT DISULFID 142 189
FT DISULFID 146 191
FT DISULFID 136 136
FT DISULFID 145 145
FT CARBOHYD 155 155
FT CARBOHYD 185 185
FT CARBOHYD 287 287
SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

Query Match 75.0%; Score 54; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC 035251;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
DE growth factor) (FltF).
GN RGF OR VEGFD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley.
RA Yamada Y., Hirata Y., Nezu J., Shlman M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
CC and endothelial cell growth, stimulating their proliferation and
CC migration and also has effects on the permeability of blood
CC vessels. May function in the formation of the venous and lymphatic
CC vascular systems during embryogenesis, and also in the maintenance
CC of differentiated lymphatic endothelium in adults. Binds and
CC activates VEGFR-3 (Flt4) receptor (By similarity).
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AF014827; AAB6557.1; -
DR HSSP: P15692; IVP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 326
FT DOMAIN 227 317
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 306 317
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 326 AA; 37112 MW; 1261FA373596C00 CRC64;

Query Match 70.8%; Score 51; DB 1; Length 326;

Best Local Similarity	90.9%	Pred. No. 0.049	Matches* 10	Conservative 1	Mismatches 0	Indels 0	Gaps 0
OY	2	ASELGRSTNTE 12	:				
Db	126	ASELGKTTNTE 136					
RESULT	3						
VEGD_MOUSE	STANDARD:	PRT:	358	AA.			
AC	P97946:						
DT	01-MAR-2002 (Rel. 41, Created)						
DT	01-MAR-2002 (Rel. 41, Last sequence update)						
DT	01-MAR-2002 (Rel. 41, Last annotation update)						
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FgfF).						
DE	FGF or VEGFD.						
GN	Mus musculus (Mouse).						
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.						
OX	NCBI_TaxID=10090;						
RP	SEQUENCE FROM N.A.						
RP	STRAIN=C57BL/6J; TISSUE=Fibroblast;						
RX	MEDLINE=97030254; PubMed=8876195;						
RA	Oriandini M., Marconcini L., Ferruzzi R., Oliviero S.;						
RT	"Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family.";						
RL	Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).						
RN	[2]						
RC	SEQUENCE FROM N.A.						
RC	TISSUE=Lung;						
RA	MEDLINE=97349118; PubMed=9205122;						
RA	Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;						
RT	"Molecular cloning of a novel vascular endothelial growth factor, VEGF-D.";						
RL	Genomics 42:483-488(1997).						
RN	[3]						
RP	DEVELOPMENTAL STAGE.						
RP	MEDLINE=98288130; PubMed=9622638;						
RA	Avantaggiato V., Oriandini M., Acampora D., Oliviero S., Simeone A.;						
RT	"Embryonic expression pattern of the murine fgf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family.";						
RL	Mech. Dev. 73:221-224(1998).						
RN	[4]						
RP	RECEPTOR SPECIFICITY.						
RP	MEDLINE=21276411; PubMed=11279005;						
RA	Baldwin K.E., Catimel B., Nice E.C., Roufail S., Hall N.E.;						
RA	Stewiers K.L., Karkkainen M.J., Allitalo K., Stacker S.A., Achen M.G.;						
RT	"The specificity of receptor binding by vascular endothelial growth factor-d is different in mouse and man.";						
RL	J. Biol. Chem. 276:19166-19171(2001).						
RC	-I- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor.						
CC	-I- SUBUNIT: Homodimer; non-covalent and antiparallel.						
CC	-I- SUPRACELLULAR LOCATION: Secreted.						
CC	-I- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.						
CC	-I- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several body structures and organs of the embryo such as limb buds, acoustic ganglion, teeth, heart, anterior pituitary as well as lung and kidney mesenchyme, liver, derma, and peritoneum of the vertebral column.						
CC	-I- INDUCTION: By the transcription factor c-fos.						
CC	-I- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward						

```

CC VEGF-3 and VEGF-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89572; CAA67892.1; -
CC EMBL; D89628; BAA14002.1; -
CC DR HSSP; P15692; IVP.
CC MGD; MGI:108037; Figf.
CC DR InterPro: IPR000072; PDGF.
CC DR Pfam: PF00341; PDGF; 1.
CC DR ProDom: PD001629; PDGF; 1.
CC DR SMART; SM00441; PDGF; 1.
CC DR PROSITE; PS00248; PDGF_1; 1.
CC DR PROSITE; PS50278; PDGF_2; 1.
CC KM Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
CC Cleavage on pair of basic residues; Multigene family.
CC STIGAL 1 21
CC PROPEP 22 93 POTENTIAL.
CC CHAIN 94 210 VASCULAR ENDOTHELIAL GROWTH FACTOR D.
CC PROPEP 211 358 POTENTIAL.
CC DOMAIN 227 323 4 X 16 AA REPEATS OF C-X(10)-C-X-C-
CC X(1,3)-C.
CC 1 (APPROXIMATE).
CC FT REPEAT 227 242 2.
CC FT REPEAT 263 278 2.
CC FT REPEAT 282 298 3.
CC FT REPEAT 306 323 4.
CC FT DISULFID 116 158 INTRACHAIN (BY SIMILARITY).
CC FT DISULFID 147 194 INTRACHAIN (BY SIMILARITY).
CC FT DISULFID 151 196 INTRACHAIN (BY SIMILARITY).
CC FT DISULFID 141 141 INTERCHAIN (BY SIMILARITY).
CC FT DISULFID 150 150 INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 358 AA; 40908 MW; 6636B17BF07037C CRC64;
SQ

```

RT "Amino acid sequence of an anti-tumor protein from Rana pipiens
 RL oocytes and early embryos. Homology to pancreatic ribonucleases.";
 RN J. Biol. Chem. 266:245-251(1991).
 [2]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-93066156; Pubmed-1438177;
 RA Mosmann S.C., Johns K.L., Ardelt W., Mikulski S.M., Shogen K.,
 RT "Comparative molecular modeling and crystallization of P-30 protein:
 a novel antitumor protein of Rana pipiens oocytes and early
 RL embryos.";
 RN Proteins 14:392-400(1992).
 [3]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RX MEDLINE-94166079; Pubmed-8120892;
 RA Mosmann S.C., Ardelt W., James M.N.G.;
 RT Refined 1.7 A X-ray crystallographic structure of P-30 protein, an
 RL amphibian ribonuclease with anti-tumor activity.";
 RN J. Mol. Biol. 236:1141-1153(1994).
 CC -1- FUNCTION: BASIC PROTEIN WITH ANTIPROLIFERATIVE/CYTOTOXIC ACTIVITY
 IN VIVO. IT EXHIBITS A RIBONUCLEASE-LIKE ACTIVITY AGAINST HIGH
 MOLECULAR WEIGHT RIBOSOMAL RNA.
 CC -1- DEVELOPMENTAL STAGE: EARLY EMBRYOS (UP TO FOUR BLASTOMERE STAGE).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 DR PDB: 1ONC: 31-JAN-94.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA: 1.
 DR ProDom: PD000535; RNaseA: 1.
 DR SMART: SM00092; RNase_Pc: 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC: 1.
 KW Hydrolyase; Nuclease; Endonuclease; 3D-structure.
 FT MOD_RES 1 1
 FT ACT_SITE 10 10
 FT ACT_SITE 31 31
 FT ACT_SITE 97 97
 FT DISULFID 19 68
 FT DISULFID 30 75
 FT DISULFID 48 90
 FT DISULFID 87 104
 FT HELIX 3 10
 FT STRAND 11 12
 FT STRAND 19 22
 FT TURN 23 24
 FT TURN 26 30
 FT STRAND 33 38
 FT HELIX 41 48
 FT TURN 49 50
 FT STRAND 55 58
 FT STRAND 63 70
 FT TURN 74 75
 FT STRAND 77 84
 FT STRAND 86 91
 FT TURN 92 93
 FT STRAND 94 101
 SQ SEQUENCE 104 AA; 11845 MW; 22A753C2F9E566B4 CRC64;

Query Match 59.7%; Score 43; DB 1; Length 104;
 Best Local Similarity 61.5%; Pred. No. 0.47;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGRKSTNTEC 13
 DB 75 CKYKLKSTNTEC 87

RESULT 5
 YNVS_CAEEL
 ID YNVS_CAEEL STANDARD; PRT; 409 AA.
 AC P34568;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothesized 46.4 kDa protein T16H12.5 in chromosome III.
 GN T16H12.5
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Smith A.;
 RT Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO C.ELEGANS F37A4.9 AND HUMAN SPO.
 CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z30662; CA83138.1;
 DR PIR: S42384; S42384.
 DR WormPep: T16H12.5; CE00510.
 DR InterPro: IPR000210; BTB_POZ.
 DR InterPro: IPR002083; MATH.
 DR Pfam: PF00651; BTB; 1.
 DR Pfam: PF00917; MATH; 1.
 DR SMART: SM00225; BTB; 1.
 DR SMART: SM00061; MATH; 1.
 DR PROSITE: PS50097; BTB; 1.
 KW Hypothetical protein.
 FT DOMAIN 223 296
 FT SEQUENCE 409 AA; 46414 MW; 7A6E116683EB098A CRC64;

Query Match 54.2%; Score 39; DB 1; Length 409;
 Best Local Similarity 58.3%; Pred. No. 9.9;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGRKSTNTE 12
 DB 229 CKSDLGSPQTQF 240

RESULT 6
 LYT8_BUCAI
 ID LYT8_BUCAI STANDARD; PRT; 319 AA.
 AC P57247;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE LYT8 protein homolog.
 GN LYT8 OR BU147
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_Taxid=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TOKYO 1998;
 RX MEDLINE-20445173; Pubmed-10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RL Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- PATHWAY: Deoxyxylulose-5-phosphate pathway (DXP) of isoprenoid
 CC biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE LYT8 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

KW Hydroxylase: Serine esterase; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 433 PECTINESTERASE B.
 FT LIPID 22 22 N-ACYL DIGLYCERIDE.
 SO SEQUENCE 433 AA: 46793 MW: 4854AD25F619B18 CRC64;

Query Match 51.4%; Score 37; DB 1; Length 433;
 Best Local Similarity 46.2%; Pred. No. 25;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGTSTNTRC 13
 Db 192 CATKAGATINTTC 204

RESULT 9
 ID PKCL_SCHPO STANDARD: PRT; 988 AA.
 AC P36582;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Protein kinase C-like 1 (EC 2.7.1.-)
 GN PKC1 OR SPAC2H10.01C OR SPAC1768.14C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93259141; Pubmed=8491190;
 RA Toda T., Shimanuki M., Yanagida M.;
 RT "Two novel protein kinase C-related genes of fission yeast are
 essential for cell viability and implicated in cell shape control.",
 RL EMBO J. 12:1987-1995(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Devlin K., Churche C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE CONTROL OF THE CELL SHAPE. TARGET
 OF THE INHIBITOR STAUFOSPORINE.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D14337; BAA03267.1;
 DR EMBL: 269730; CAA93602.1;
 DR EMBL: 269795; CAA93697.1;
 DR PIR: S35362; S35362.
 DR HSSP: Q63450; IAD6.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR000861; RKM_repeat.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00130; DAG_PE-bind; 2.
 DR Pfam: PF02185; HRI; 2.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00433; pkinase_C; 1.
 DR SMART: SM00109; C1; 2.

DR SMART: SM00239; C2; 1.
 DR SMART: SM00074; HRI; 2.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TK; 1.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
 KM Phorbol-ester binding; Repeat.
 FT DOMAIN 414 461 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 481 530 PHORBOL-ESTER AND DAG BINDING 2.
 FT NP_BIND 664 923 PROTEIN KINASE.
 FT NP_BIND 670 678 ATP (BY SIMILARITY).
 FT BINDING 693 693 ATP (BY SIMILARITY).
 FT ACT_SITE 789 789 BY SIMILARITY.
 FT CONFLICT 27 35 AMVASTKRP -> SNGGRDGS (IN REF. 1).
 SO SEQUENCE 988 AA: 111783 MW: 0959BDECIAB43C4E CRC64;

Query Match 51.4%; Score 37; DB 1; Length 988;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 LGKSTNTRC 13
 Db 817 LGMTSTTC 825

RESULT 10
 ID KPCL_COCHE STANDARD: PRT; 1174 AA.
 AC 042632;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein kinase C-like (EC 2.7.1.-).
 GN PKC1.
 OS Cochlidiobolus heterostrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochlidiobolus.
 OX NCBI_TaxID=5016;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2 / ATCC 48329;
 RA Oeser B.M., Yoder O.C.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y15839; CAA75801.1;
 DR HSSP: P05132; IADP.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR000861; RKM_repeat.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00130; DAG_PE-bind; 2.
 DR Pfam: PF02185; HRI; 2.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00433; pkinase_C; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X74791; CAA52796.1; -.
 DR PIR: S37055; S37055.
 DR HSSP: P42321; 2CAE.
 DR InterPro: IPR002226; Catalase.
 DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PR00067; CATALASE.
 DR ProDom: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE_1; 1.
 DR PROSITE: PS00438; CATALASE_2; 1.
 KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
 FT ACT_SITE 54 54 BY SIMILARITY.
 FT ACT_SITE 127 127 BY SIMILARITY.
 FT BINDING 337 337 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 483 AA; 54087 MW; B58CF8230B8A2F55 CRC64;

Query Match 50.0%; Score 36; DB 1; Length 483;
 Best Local Similarity 70.0%; Pred. No. 42;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 3 SPLGKSTNTE 12
 ||:|| | ||
 Db 80 SEVGKRTETP 89

Search completed: June 17, 2002, 16:24:25
 Job time: 1554 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:46 ; Search time 172.85 seconds
(without alignments)
13.011 Million cell updates/sec

Title:	US-09-761-636A-5
Perfect score:	72
Sequence:	1 CASELGKSTNTFC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :

1:  sp_Archaea:*
2:  sp_Bacteria:*
3:  sp_Fungi:*
4:  sp_Human:*
5:  sp_Invertebrate:*
6:  sp_Mammal:*
7:  sp_mhc:*
8:  sp_Organelle:*
9:  sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_Vertebrate:*
14: sp_unclassified:
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	70.8	326	11	Q91ZE4	Q91ZE4 ratlvs norv
2	49	68.1	127	13	Q918V8	Q918V8 rana pipien
3	43	59.7	997	5	Q967Z1	Q967Z1 trypanosoma
4	42	58.3	1011	3	Q9P944	Q9P944 pneumocysti
5	41	56.9	316	13	Q9QZB2	Q9QZB2 brachydanio
6	41	56.9	397	16	Q9ZCN9	Q9ZCN9 listeria in
7	40	55.6	383	12	Q84551	Q84551 paramacium
8	39	54.2	128	13	Q9DFV7	Q9dfv7 rana catesb
9	39	54.2	471	10	Q9FHK1	Q9fHK1 arabidopsis
10	39	54.2	677	10	Q9FLD9	Q9fLD9 arabidopsis
11	39	54.2	1117	10	Q9M133	Q9m133 arabidopsis
12	38	52.8	23	2	Q9R4A0	Q9r4A0 lactobacill
13	38	52.8	24	3	Q9R499	Q9r499 lactobacill
14	38	52.8	49	3	Q9U0P2	Q9u0P2 pneumocysti
15	38	52.8	188	13	Q9DFY8	Q9dfY8 rana catesb
16	38	52.8	254	15	Q85498	Q85498 human immun

17	38	52.8	254	15	Q9YX6	Q9YX6 human	immun
18	38	52.8	255	15	O89503	O89503 human	immun
19	38	52.8	258	3	P87104	P87104 pneumocysti	
20	38	52.8	258	2	Q9S332	Q9S332 prochloroco	
21	38	52.8	334	15	Q76114	Q76114 human	immun
22	38	52.8	334	15	Q76115	Q76115 human	immun
23	38	52.8	334	15	Q76116	Q76116 human	immun
24	38	52.8	334	15	Q76080	Q76080 human	immun
25	38	52.8	334	15	Q76083	Q76083 human	immun
26	38	52.8	334	15	Q76084	Q76084 human	immun
27	38	52.8	334	15	Q76088	Q76094 human	immun
28	38	52.8	334	15	Q76094	Q76098 human	immun
29	38	52.8	334	15	Q76098	Q76101 human	immun
30	38	52.8	334	15	Q76101	Q76104 human	immun
31	38	52.8	334	15	Q76104	Q76106 human	immun
32	38	52.8	334	15	Q76106	Q76107 human	immun
33	38	52.8	334	15	Q76107	Q76108 human	immun
34	38	52.8	334	15	Q76108	Q76109 human	immun
35	38	52.8	334	15	Q76109	Q76110 human	immun
36	38	52.8	334	15	Q76110	Q76113 human	immun
37	38	52.8	334	15	Q76113	Q97GW2 clostridium	
38	38	52.8	378	16	Q97GW2	Q97GW2 clostridium	
39	38	52.8	393	15	Q41549	Q41549 human	immun
40	38	52.8	399	15	Q41551	Q41551 human	immun
41	38	52.8	413	15	Q41550	Q41550 human	immun
42	38	52.8	418	15	Q41548	Q41548 human	immun
43	38	52.8	469	10	O9SCK6	O9SCK6 arabidopsis	
44	38	52.8	520	3	P87105	P87105 pneumocysti	
45	38	52.8	552	1	O51938	O51938 methanosarc	

ALIGNMENTS

RESULT	1	
Q91ZE4		
ID	Q91ZE4	PRELIMINARY;
AC	Q91ZE4;	PRT: 326 AA.
DT	01-DEC-2001	(TREMBLrel. 19, Created)
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE	VEGF-D.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCHI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SPRAGUE-DAWLEY;	
RA	Kirklin V., Matitschek R., Striffler J., Steffen A., Waltenberger J.,	
RA	Pepper M.S., Giannis A., Sleeman J.P.;	
RT	"Characterization of indolones which specifically inhibit VEGF-C and	
RT	VEGF-D-induced activation of VEGFR-3 but not VEGFR-2.";	
RL	Eur. J. Biochem. 0:0-0(2001).	
EMBL:	AY032728; AAK96008.1; -	
SO	SEQUENCE 326 AA; 37106 MW; D7CAEBA6C9FABB7D CRC64;	

Query Match	70.8%;	Score 51;	DB 11;	Length 326;
Best Local Similarity	90.9%;	Pred. No. 0.18;		
Matches	10;	Conservative	1;	Mismatches 0; Gaps 0.
QY	2	ASELGKSTNTF	12	
		:		
DB	126	ASELGKSTNTF	136	
RESULT	2			
Q918V8				
ID	Q918V8	PRELIMINARY;	PRT;	127 AA.
AC	Q918V8;			
DC	01-OCT-2000 (TREMBLrel, 15, Created)			
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)			

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE ONCOMASE VARIANT RAPLRI PRECURSOR.
 OS Rana pipiens (Northern Leopard frog).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 CC Amphibia: Batrachia: Anura: Neobatrachia: Ranioidea: Rana.
 OX NCBI_TaxId=8404;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=20330357; PubMed=10871370;
 RA Chen S.-L., Le S.-Y., Newton D.L., Maizel J.V. Jr., Rybak S.M.;
 RT "A gender-specific mRNA encoding a cytotoxic ribonuclease contains a
 RT 3' UTR of unusual length and structure."
 RL Nucleic Acids Res. 28:2375-2382(2000).
 DR EMBL: AF165133; AAF76935.1; -
 DR HSSP: P22069; IONC.
 DR InterPro: IPR001427; RNaseH.
 DR Pfam: PF00074; RNaseA; 1.
 DR ProDom: PD000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL
 SQ SEQUENCE 127 AA; 14491 MW; B851DC5407AB69B CRC64;

Query Match 68.1%; Score 49; DB 13; Length 127;
 Best Local Similarity 69.2%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CASELGKSTNFC 13
 | : | | | | | | | | | |
 Db 98 CKYKLRKSTNFC 110

RESULT 3
 ID 096721 PRELIMINARY; PRT; 997 AA.
 AC 096721;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DE PUTATIVE MISMATCH REPAIR PROTEIN MSH8.
 OS Trypanosoma brucei.
 OC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Trypanosoma.
 OX NCBI_TaxId=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MITA1.2;
 RA Bell J.S., Harvey T.I., Sims A.M., Barry J.D., McCulloch R.;
 RT "Cloning and characterization of two Muts homologs and two MutL
 RT homologs from Trypanosoma brucei."
 RL Submitted (FE8-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF350880; AAK51796.1; -
 SQ SEQUENCE 997 AA; 110417 MW; DBFDCA3D8D2642A3 CRC64;

Query Match 59.7%; Score 43; DB 5; Length 997;
 Best Local Similarity 58.3%; Pred. No. 18;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CASELGKSTNFC 12
 | : | : | | | | | | | | | |
 Db 460 GSKRGRTNTE 471

RESULT 4
 ID 09P944 PRELIMINARY; PRT; 1011 AA.
 AC 09P944;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)

DE KEXIN-LIKE PROTEASE KEX1.
 GN KEX1.
 OS Pneumocystis carinii f. sp. muris.
 OC Eukaryota: Fungi: Ascomycota: Pneumocystidomycetes: Pneumocystidaceae;
 CC Pneumocystis.
 OX NCBI_TaxId=42066;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20184731; PubMed=10721706;
 RA Lee L.H., Gigliotti F., Wright T.W., Simpson-Haldaris P.J.,
 RA Weinberg G.A., Haldaris C.G.;
 RT "Molecular characterization of KEX1, a kexin-like protease in mouse
 RT Pneumocystis carinii."
 RL Gene 242:141-150(2000).
 DR EMBL: AF093132; AAF32493.1; -
 DR MEROPS: S08.011; -
 DR InterPro: IPR00209; Peptidase_S8.
 DR InterPro: IPR002884; P_domain.
 DR InterPro: IPR002965; P_rich_extensn.
 DR Pfam: PF01483; P; 1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR01217; PRICHEXTENS.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000717; P_domain; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 KW Protease.
 SQ SEQUENCE 1011 AA; 112021 MW; FBE472C8F65864E8 CRC64;

Query Match 58.3%; Score 42; DB 3; Length 1011;
 Best Local Similarity 53.8%; Pred. No. 28;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CASELGKSTNFC 13
 | | : : | | | | | | | | | |
 Db 246 CAGEIVAKNTFC 258

RESULT 5
 ID 090ZB2 PRELIMINARY; PRT; 316 AA.
 AC 090ZB2;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DE TRANSCRIPTION FACTOR GBX1.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
 CC Actinopterygii: Neopterygii: Teleostei: Euteleostei: Ostariophysi;
 OC Cypriniformes: Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lun K., Thisse C., Thisse B., Amores A., Yan Y., Postlethwait J.,
 RA Brand M.;
 RT "Gbx2 but not gbx1 is dependent on fgf8 function."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF288763; AAK63070.1; -
 SQ SEQUENCE 316 AA; 34408 MW; DC32955EA5430D8A CRC64;

Query Match 56.9%; Score 41; DB 13; Length 316;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ASELGKSTNFC 13
 | | : | : | | | | | | | | | |
 Db 63 ASFAGRLTNFC 74
 RESULT 6
 Q92C9N

ID Q92CN9 PRELIMINARY; PRT; 397 AA.
 AC Q92CN9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE ACKA2 PROTEIN.
 GN ACKA2 OR LIN1132.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed-11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 Madeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Bojand J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 DR EMBL:AL596167; CAC96363.1; -;
 DR Listlist; LIN01132; -;
 KW Complete proteome.
 SQ SEQUENCE 397 AA; 43115 MW; 5822544EF92CBF51 CRC64;

Query Match 56.9%; Score 41; DB 16; Length 397;
 Best Local Similarity 72.7%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CASELCKSTNT 11
 11 111111
 Db 214 CALCKSVNT 224

RESULT 7
 ID Q84551 PRELIMINARY; PRT; 383 AA.
 AC Q84551;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE A231L PROTEIN.
 GN A231L.
 OS Paramesecium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96187795; PubMed-8614977;
 RA Lu Z., Li Y., Que O., Kutish G.F., Rock D.L., Van Etten J.L.;
 RT "Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map
 positions 88 to 182.";
 RL Virology 216:102-123(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20013326; PubMed-10544099;
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
 RT Lisee A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homosperridine
 synthase.";
 RL Virology 263:254-262(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20478054; PubMed-11021991;

RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 PBCV-1.";
 RL Virology 276:27-36(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL:U42580; AAC96599.1; -;
 SQ SEQUENCE 383 AA; 43643 MW; B04E602D99973DCE CRC64;

Query Match 55.6%; Score 40; DB 12; Length 383;
 Best Local Similarity 46.2%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 CASELCKSTNTFC 13
 1 11 11
 Db 304 CWMICKKKNKFC 316

RESULT 8
 ID Q9DFY7 PRELIMINARY; PRT; 128 AA.
 AC Q9DFY7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-OCT-2001 (TREMBLrel. 19, Last sequence update)
 DE RC-RNASE3 RIBONUCLEASE PRECURSOR.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE-20512555; PubMed-11058105;
 RA Liao Y.D., Huang H.C., Leu Y.J., Wei C.W., Tang P.C., Wang S.C.;
 RT "Purification and cloning of cytototoxic ribonucleases from Rana
 catesbeiana (bullfrog).";
 RL Nucleic Acids Res. 28:4097-4104(2000).
 DR EMBL:AF242554; AAG31440.2; -;
 DR HSP: P22069; IONC.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR ProDom: PD000535; RNaseA; 1.
 DR SMART: SM0092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL. 1 23 POTENTIAL.
 FT CHAIN 24 128 RC-RNASE3 RIBONUCLEASE.
 FT FT 128
 SQ SEQUENCE 128 AA; 14517 MW; 2B14986082E0587D CRC64;

Query Match 54.2%; Score 39; DB 13; Length 128;
 Best Local Similarity 53.8%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASELGKSTNTPC 13
 DB 98 CHYKLSSTNTIC 110

RESULT 9
 Q9FHK1 PRELIMINARY; PRT; 471 AA.

AC Q9FHK1; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 OS GENOMIC DNA, CHROMOSOME 5, PI CLONE:MAB16.
 OC Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 DR EMBL; AB018112; BAB10975.1;
 DR EMBL; AB026661; BAB10975.1; JOINED.
 DR InterPro: IPR001810; F-box.
 DR Pfam: PF00646; F-box; 1.
 DR SMART; SM00256; FBOX; 1.
 DR PROSITE; PS50181; FBOX; 1.
 SQ SEQUENCE 471 AA; 55212 MW; 1E5D997B7942E26C CRC64;

Query Match 54.2%; Score 39; DB 10; Length 471;
 Best Local Similarity 60.0%; Pred. No. 47;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELGKSTNTPC 13
 DB 365 DLGKSTNTKC 374

RESULT 10
 Q9FLD9 PRELIMINARY; PRT; 677 AA.

AC Q9FLD9; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 OS SIMILARITY TO HEDGEGHOG-INTERACTING PROTEIN.
 OC Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98290546; PubMed=9628582;
 RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones.";

RL DNA Res. 5:41-54(1998).
 DR EMBL; AB010077; BAB10221.1;
 SQ SEQUENCE 677 AA; 74196 MW; F0C2B09BB6E1B079 CRC64;

Query Match 54.2%; Score 39; DB 10; Length 677;
 Best Local Similarity 63.6%; Pred. No. 68;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 SELGKSTNTPC 13
 DB 144 SEIKSSNDPC 154

RESULT 11
 Q9M133 PRELIMINARY; PRT; 1117 AA.

AC Q9M133; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 OS HYPOTHETICAL 126.0 KDA PROTEIN.
 GN ATG601400.
 OC Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL161491; CAB80949.1;
 DR InterPro: IPR002885; PPR.
 DR Pfam: PF01535; PPR; 4.
 KW Hypothetical protein.
 SQ SEQUENCE 1117 AA; 126000 MW; 775D642AB6D83AB6 CRC64;

Query Match 54.2%; Score 39; DB 10; Length 1117;
 Best Local Similarity 58.3%; Pred. No. 11e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CASELGKSTNTP 12
 DB 922 CLSELGELSTP 933

RESULT 12
 Q9R4A0 PRELIMINARY; PRT; 23 AA.

AC Q9R4A0; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DE 01-MAY-2000 (TREMBlrel. 13, last annotation update)
 OS ACIDOCIN J1132 ALPHA PEPTIDE (FRAGMENT).
 OC Lactobacillus acidophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Lactobacillus.
 NCBI_TaxID=1579;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=97076919; PubMed=8975617;
 RA Tahara T., Oshimura M., Umezawa C., Kanatani K.;
 RT "Isolation, partial characterization, and mode of action of Acidocin
 RT J1132, a two-component bacteriocin produced by Lactobacillus
 RT acidophilus JCM 1132.";
 RL Appl. Environ. Microbiol. 62:892-897(1996).

SO SEQUENCE 23 AA; 2269 MW; 45231D87B690772E CRC64;

Query Match 52.8%; Score 38; DB 2; Length 23;
Best Local Similarity 66.7%; Pred. No. 3.2;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKST 9
DB 7 CASQIGKST 15

RESULT 13

O9R499 PRELIMINARY; PRT; 24 AA.
AC O9R499;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE ACIDOCIN J1132 BETA PEPTIDE (FRAGMENT).
OC Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]
RP MEDLINE=97076919; PubMed=8975617;
RA Tahara T., Oshimura M., Umezawa C., Kanatani K.;
RT "Isolation, Partial Characterization, and mode of action of Acidocin
RT J1132, a two-component bacteriocin produced by Lactobacillus
RT acidophilus JCM 1132."
RL Appl. Environ. Microbiol. 62:892-897(1996).
SQ SEQUENCE 24 AA; 2326 MW; E1141D87B690775D CRC64;

Query Match 52.8%; Score 38; DB 2; Length 24;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKST 9
DB 8 CASQIGKST 16

RESULT 14

O9UUP2 PRELIMINARY; PRT; 49 AA.
AC O9UUP2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KEXIN (FRAGMENT).
OC Pneumocystis carinii f. sp. carinii.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=38081;
RN [1]
RP SEQUENCE FROM N.A.
RA Russian D.A., Andrawis-Sorail V., Kovacs J.A.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U82985; AAF21236.1; -;
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
FT NON_TER 1 49
FT 1 49
SQ SEQUENCE 49 AA; 5205 MW; 326DAF0626AC0420 CRC64;

Query Match 52.8%; Score 38; DB 3; Length 49;
Best Local Similarity 46.2%; Pred. No. 7;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 CASELGKSTNTPC 13

DB 28 CAGEVAAARNITLC 40

RESULT 15

O9DFY8 PRELIMINARY; PRT; 128 AA.
ID O9DFY8;
AC O9DFY8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RC-RNASE2 RIBONUCLEASE PRECURSOR.
OC Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=20512555; PubMed=11058105;
RA Liao Y.-D., Huang H.-C., Liao Y.-J., Wei C.-W., Tang P.-C.,
RT "Purification and cloning of cytotoxic ribonucleases from Rana
RT catesbeiana (bullfrog)."
RL Nucleic Acids Res. 28:4097-4104(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Liao Y.-D., Huang H.-C., Liao Y.-J., Wei C.-W., Tang P.-C.,
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF242533; AAG31439.1; -;
DR HSSP; P22069; IONC.
DR InterPro: IPR001427; RNaseA.
DR Pfam: PF00074; RNaseA; 1.
DR Prodom: PD000535; RNaseA; 1.
DR SMART: SM00092; RNase-PC; 1.
DR PROSITE: PS00127; RNase_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 23
FT 1 23
FT CHAIN 24 128
SQ SEQUENCE 128 AA; 14839 MW; 989719CF52053BCC CRC64;

Query Match 52.8%; Score 38; DB 13; Length 128;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGKSTNTPC 13
DB 98 CHYKTKSSNTIC 110

Search completed: June 17, 2002, 16:23:34
Job time: 1608 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:55:11 ; Search time 227.35 Seconds
(without alignments)
6.351 Million cell updates/sec

Title: US-09-761-636A-5
Perfect score: 72
Sequence: 1 CASLGKSTNTPC 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: A-Genesec 032802.*
2: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	72	100.0	13	22	AAU04524	VEGF based monocyc
2	72	100.0	17	22	AAU04527	VEGF based bicycl1
3	69	95.8	13	22	AAU04534	VEGF based monocyc
4	69	95.8	17	22	AAU04546	VEGF based bicycl1
5	69	95.8	17	22	AAU04550	VEGF based bicycl1
6	65	90.3	13	22	AAU04535	VEGF based monocyc
7	65	90.3	13	22	AAU04537	VEGF based monocyc
8	65	90.3	17	22	AAU04547	VEGF based bicycl1
9	65	90.3	17	22	AAU04549	VEGF based bicycl1
10	65	90.3	17	22	AAU04551	VEGF based bicycl1
11	61	84.7	13	22	AAU04536	VEGF based monocyc

12	61	84.7	17	22	AAU04548	VEGF based bicycl1
13	54	75.0	73	22	AAU04522	Human VEGF-D amino
14	54	75.0	96	22	AAU04520	Human VEGF-D amino
15	54	75.0	109	20	AAV23889	Human vascular end
16	54	75.0	109	21	AAV11931	Human truncated VE
17	54	75.0	325	19	AAW53240	Homo sapiens vascu
18	54	75.0	325	22	AAV97572	Human VEGF-D prote
19	54	75.0	354	19	AAW49036	Homo sapiens vascu
20	54	75.0	354	19	AAW53241	Human vascular end
21	54	75.0	354	19	AAW44293	Human VEGF-D prote
22	54	75.0	354	21	AAV10649	Human VEGF-D prote
23	54	75.0	354	21	AAV29049	Human VEGF-D prote
24	54	75.0	354	21	AAV70750	Human VEGF-D prote
25	54	75.0	354	21	AAV70983	Human VEGF-D prote
26	54	75.0	354	22	AAU08441	Human VEGF-D prote
27	54	75.0	354	22	AAV70685	Human VEGF-D prote
28	54	75.0	354	22	AAV97573	Human VEGF-D prote
29	54	75.0	354	22	AAV37606	Human VEGF-D prote
30	54	75.0	620	18	AAV14594	Human growth facto
31	54	75.0	178	20	AAV08287	Human growth facto
32	51	70.8	321	19	AAW53243	Human vascular end
33	51	70.8	326	19	AAW44296	Human vascular end
34	51	70.8	337	20	AAV08286	Human growth facto
35	51	70.8	358	18	AAV14592	Human growth facto
36	51	70.8	358	19	AAV53242	Human vascular end
37	51	70.8	358	19	AAV42425	Human vascular end
38	49	68.1	104	20	AAV28865	Human growth facto
39	49	68.1	104	20	AAV28866	Human growth facto
40	49	68.1	104	20	AAV28867	Human growth facto
41	49	68.1	105	20	AAV28869	Human growth facto
42	49	68.1	105	20	AAV28870	Human growth facto
43	49	68.1	105	20	AAV28871	Human growth facto
44	49	68.1	127	20	AAV28872	Human growth facto
45	43	59.7	83	18	AAV53119	Human growth facto

ALIGNMENTS

RESULT 1
AAU04524
ID AAU04524 standard; Peptide: 13 AA.

AC AAU04524;

DT 26-SEP-2001 (first entry)

DE VEGF based monocyclic peptide 1.

Human: VEGF: vascular endothelial growth factor; angiogenesis;
neovascularisation; lymphangiogenesis; psoriasis; tumour;
diabetic induced neovascular sequelae; rheumatoid arthritis;
diabetic retinopathy; chronic inflammation; cyclic.

Synthetic.

Key Disulfide-bond 1.13 location/Qualifiers
/note="This bond cyclises the peptide"

WO200152875-A1.
26-JUL-2001.

18-JAN-2001; 2001WO-US01533.

18-JAN-2000; 2000US-0176293.

16-MAY-2000; 2000US-0204590.

(LUDWIG-) LUDWIG INST CANCER RES.

Achen MG, Hughes RA, Stacker S, Gendron A;

XX

DR WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -

PS Claim 49; Page 32; 102pp; English.

CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC deleted prior to cyclisation) and a cyclic peptide with at least one amino acid
 CC neovascularisation or lymphangiogenesis in a mammal with angiogenesis,
 CC characterised by diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heart or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

SQ Sequence 13 AA;

Query Match 100.0%; Score 72; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASELKGSTNTFC 13
 Db 1 caseLgkstntfc 13

RESULT 2
 AAU04527
 ID AAU04527 standard; Protein; 17 AA.
 XX
 AC AAU04527;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based bicyclic dimeric peptide #1.
 XX
 KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key
 FT Disulfide-bond 1.13 Location/Qualifiers
 FT Disulfide-bond 17 /note= "This bond cyclises the peptide"
 FT /note= "A disulfide bond forms between residue
 FT 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide, or to residue 1 of the sequence

FT appearing as AAU04528, also forming a dimeric peptide"

XX WO200152875-A1.

PN 26-JUL-2001.

PD 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

XX (LUDM-) LUDMIG INST CANCER RES.

PA Achen MG, Hughes RA, Stackers S, Cendron A;
 PI WPI: 2001-442248/47.

DR Novel monomeric monocyclic peptide, used to interfere with
 XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 XX loop fragment from an exposed loop of a growth factor protein by
 XX oxidizing the cysteine residues -

PS Claim 59; Page 32; 102pp; English.

CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC deleted prior to cyclisation) and a cyclic peptide with at least one amino acid
 CC neovascularisation or lymphangiogenesis in a mammal with angiogenesis,
 CC characterised by diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heart or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

SQ Sequence 17 AA;

Query Match 100.0%; Score 72; DB 22; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASELKGSTNTFC 13
 Db 1 caseLgkstntfc 13

RESULT 3
 AAU04534
 ID AAU04534 standard; Peptide; 13 AA.
 XX
 AC AAU04534;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based monocyclic peptide 12.

XX Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX Synthetic.

XX Key Location/Qualifiers
 FT Disulfide-bond 1..13
 FT /note= "this bond cyclises the peptide"

XX WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

XX 16-MAY-2000; 2000US-0204590.

XX (LUDM-) LUDMIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;
 WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX Example 25; Page 47; 102pp; English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

XX Sequence 13 AA;

Query Match 95.8%; Score 69; DB 22; Length 13;
 Best local Similarity 92.3%; Pred. No. 3.3e-05;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGRSTNTEC 13
 DB 1 catelgrstntfc 13

RESULT 4

AAU04546 standard; Peptide; 17 AA.

AAU04546;

26-SEP-2001 (first entry)

VEGF based bicyclic dimeric peptide #3.

Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX Synthetic.

XX Key Location/Qualifiers
 FT Disulfide-bond 1..13
 FT /note= "this bond cyclises the peptide"

FT Disulfide-bond 17
 FT /note= "A disulfide bond forms between residue
 FT 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide"

XX WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

XX 16-MAY-2000; 2000US-0204590.

XX (LUDM-) LUDMIG INST CANCER RES.
 XX Achen MG, Hughes RA, Stacker S, Cendron A;
 WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX Example 26; Page 49; 102pp; English.

XX The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a

	Query Match	95.8%	Score 69;	DB 22;	Length 17;
	Best Local Similarity	92.3%	Pred. NO.	4.3e-05;	
	Matches 12;	Conservative	1;	Mismatches	0;
QY	1 CASELGGKSTNFC	13			Gaps 0;
	:				
	db 1 cateIggkstnfc	13			

RESULT	5
AAU04550	
ID	AAU04550 standard; Peptide; 17 AA.
XX	

DT 26-SEP-2001 (first entry)
XX

VEGF based bicyclic dimeric peptide #7.

KM vascular endothelial growth factor; angiogenesis
 KM neovascularisation; lymphangiologenesis; psoriasis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation; cyclic

05 Synthetic.

Key	Location/Qualifiers
FH	
EN	

FT	/note= "This bond cyclises the peptide"
FT	17
FT	

/note- "A disulfide bond forms between residue 17 and residue 1 of the sequence appearing as AAU04528, forming a dimeric peptide"

PN WO200152875-A1

PD 26-JUL-2001.

18-JAN-2001; 2001WO-US01533.

18-JAN-2000; 2000US-0176293.
16-MAY-2000: 2000US-0304500

PA (LUDW-) LUDWIG INST CANCER RES

Achen MG, Hughes RA, Stackner S, Cendron A:

WPI; 2001-442248/47.

Novel monomeric monocyclic peptide, used to interfere with angiogenesis or lymphangiogenesis.

or lymphangiogenesis, is produced by cyclising a peptide loop fragment from an exposed loop of a growth factor protein by oxidizing the cysteine residues.

Example 26; Page 49; 102pp; English.

This sequence represents a dimeric bicyclic peptide of the invention, whose 3-dimensional structure is modelled on the expose loop of human VEGF₁ (vascular endothelial growth factor). The invention relates to a method of producing a monomeric monocyclic peptide by a measuring a beta-beta carbon separation distances on opposite antiparallel strands of a peptide loop fragment from an exposed loop of a growth factor protein and cyclising the peptide by oxidising the cysteine residues. The monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked monocyclic peptides), and a cyclic peptide with at least one amino acid deleted prior to cyclisation are used to interfere with angiogenesis, neovascularisation or lymphangiogenesis in a mammal with a condition

characterised by angiogenesis, neovascularisation or lymphangiogenesis. The condition is diabetic retinopathy, psoriasis, arteriopathy, hemangioma, vascularised malignant or benign tumour, post-recovery trauma, substance-induced neovascularisation of the liver, excessive hormone-related angiogenic dysfunction, diabetes induced neovascular infection. The peptides are also used to modulate vascular permeability in a mammal (the mammal has a condition characterised by fluid accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura, or brain. The peptides are used to image blood vessels and lymphatic vasculature. The monomeric and bicyclic peptides are used to interfere with at least one biological activity induced by VEGF, VEGF-C or -D and chronic inflammation, especially rheumatoid arthritis, psoriasis and diabetic retinopathy.

Query Match	95.88;	Score 69;	DB 22;	Length 17;
Best Local Similarity	92.3%;	Pred. No. 4.3e-05;		
Matches 12; Conservative	1;	Mismatches	0;	Indels 0;

Qy	1	CASELGKSTNTEC	13
		:	
Dd	1	case1gkstnctnc	13

RESULT 6
AAU04535

XX
AC 221104535

DT 26-SEP-2001 (Fri Oct 05:00:00)

VEGF based monocyclic peptide 13.

KM human; VEGF; vascular endothelial growth factor; angiogenesis;
KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
KM diabetes induced neovascular sequelae; rheumatoid arthritis;
KM diabetic retinopathy; chronic inflammation; cyclic.

Synthetic.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT /note= "This bond cyclises the peptide"

PN WO200152875-A1
XX

26-JUL-2001.

18-JAN-2001; 2001WO-US01533.

18-JAN-2000; 2000US-0176293.
16-MAY-2000; 2000US-0204500.

(LUDW-) LUDWIG INST CANCER RES.

Achen MG, Hughes RA, Stacker S, Cendron A:

WPI; 2001-442248/47.

Novel monomeric monocyclic peptide, used to interfere with angiogenesis, or lymphangiogenesis, is produced by cyclizing a peptide loop fragment from an exposed loop of a growth factor protein by oxidizing the cysteine residues.

Example 25; Page 47; 102pp; English.

The sequence represents a monomeric monocyclic peptide of the invention,

CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

XX
 SQ Sequence 13 AA:

Query Match 90.3%; Score 65; DB 22; Length 13;
 Best Local Similarity 84.6%; Pred. No. 0.00016;
 Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKSTNTPC 13
 |||||:||||
 Db 1 caselgrstntfc 13

RESULT 7
 AA004537
 ID AA004537 standard; Peptide; 13 AA.

XX AC AA004537;

XX DT 26-SEP-2001 (first entry)

XX DE VEGF based monocyclic peptide 15.

XX KW Human: VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Disulfide-bond 1..13 /note- "this bond cyclises the peptide"

XX PN WO200152875-A1.

XX PD 26-JUL-2001.

XX PE 18-JAN-2001; 2001WO-US01533.

XX PR 18-JAN-2000; 2000US-0176293.

XX PR 16-MAY-2000; 2000US-0204590.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Achen MG, Hughes RA, Stackner S, Cendron A;

DR WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -

XX Example 25; Page 47; 102pp; English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

XX SQ Sequence 13 AA:

Query Match 90.3%; Score 65; DB 22; Length 13;
 Best Local Similarity 84.6%; Pred. No. 0.00016;
 Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKSTNTPC 13
 |||||:||||
 Db 1 caselgrstntfc 13

RESULT 8

AA004547
 ID AA004547 standard; Peptide; 17 AA.

XX AC AA004547;

XX DT 26-SEP-2001 (first entry)

XX DE VEGF based bicyclic dimeric peptide #4.

XX KW Human: VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Disulfide-bond 1..13 /note- "this bond cyclises the peptide"

XX FT Disulfide-bond 17

XX /note- "A disulfide bond forms between residue
 17 and residue 17 of an identical peptide to form
 a dimeric peptide"

```

XX XX WO200152875-A1.
XX PD 26-JUL-2001.
XX PF 18-JAN-2001; 2001WO-US01533.
XX PR 18-JAN-2000; 2000US-0176293.
XX PR 16-MAY-2000; 2000US-0204590.
XX PA (LUDMW) LUDWIG INST CANCER RES.
XX PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX DR WPI: 2001-442248/47.
XX PT Novel monomeric monocyclic peptide, used to interfere with
XX PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
XX PT loop fragment from an exposed loop of a growth factor protein by
XX PT oxidising the cysteine residues -
XX PS
XX PS Example 26; Page 49; 102pp; English.
CC CC The sequence represents a dimeric bicyclic peptide of the invention,
CC CC whose 3-dimensional structure is modelled on the expose loop of human
CC CC VEGFD (vascular endothelial growth factor). The invention relates to a
CC CC method of producing a monomeric monocyclic peptide by a measuring
CC CC a beta-beta carbon separation distances on opposite antiparallel strands of
CC CC a peptide loop fragment from an exposed loop of a growth factor protein
CC CC and cyclising the peptide by oxidising the cysteine residues. The
CC CC monocyclic peptides), dimeric bicyclic peptides (comprising 2 linked
CC CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC CC cerebrovascular accident, post-angioplasty restenosis head, heart or cold
CC CC trauma, substance-induced neovascularisation of the liver, excessive
CC CC hormone-related angiotensin dysfunction, diabetes induced neovascular
CC CC sequelae, hypertension induced neovascularisation or chronic liver
CC CC infection. The peptides are also used to modulate vascular permeability
CC CC in a mammal (the mammal has a condition characterised by fluid
CC CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC CC or brain). The peptides are used to image blood vessels and lymphatic
CC CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC CC are also used in combination with an anti-inflammatory agent, to treat a
CC CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC CC diabetic retinopathy.
XX SO Sequence 17 AA:
OY Query Match 90.3%; Score 65; DB 22; Length 17;
OY Best Local Similarity 84.6%; Pred. No. 0.00021;
DB Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 CASELGKSNTEFC 13
OY |||||:::|||
DB 1 caselgksntfc 13
RESULT 9
AAU04549
ID AU04549 standard; Peptide: 17 AA.
XX AAU04549:
XX 26-SEP-2001 (first entry)
XX VEGF based bicyclic dimeric peptide #6.

```

Query Match	Score 65:	DB 22:	Length 17:
Best Local Similarity	84.6%	Pred. No. 0.00021.	
Matches 11:	Conservative 2:	Mismatches 0:	Indels 0: Gaps 0:

Db 1 case1grstnsfc 13

```

RESULT 10
AAU04551
ID AAU04551 standard; Peptide; 17 AA.
XX
AC AAU04551:
XX
DT 26-SEP-2001 (first entry)
XX
DE VEGF based bicyclic dimeric peptide #8.
XX
KW Human: VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..13 /note= "This bond cyclises the peptide"
FT Disulfide-bond 17 /note= "A disulfide bond forms between residue
FT 17 and residue 1 of the sequence appearing as AAU04528,
FT forming a dimeric peptide"
XX
PN WO200152875-A1.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01533.
XX
PR 18-JAN-2000; 2000US-0176293.
PR 16-MAY-2000; 2000US-0204590.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX
DR WPI: 2001-442248/47.
XX
PT Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
PS Example 26; Page 49; 102pp; English.
XX
XX The sequence represents a dimeric bicyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
XX VEGFD (vascular endothelial growth factor). The invention relates to a
XX method of producing a monomeric monocyclic peptide by a measuring
XX beta-beta carbon separation distances on opposite antiparallel strands of
XX a peptide loop fragment from an exposed loop of a growth factor protein
XX and cyclising the peptide by oxidising the cysteine residues. The
XX monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX monocyclic peptides) and a cyclic peptide with at least one amino acid
XX deleted prior to cyclisation are used to interfere with angiogenesis,
XX neovascularisation or lymphangiogenesis in a mammal with a condition
XX characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX The condition is diabetic retinopathy, psoriasis, arthropathy,
XX hemangioma, vascularised malignant or benign tumour, post-recovery
XX cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX trauma, substance-induced neovascularisation of the liver, excessive
XX hormone-related angiogenic dysfunction, diabetes induced neovascular
XX sequelae, hypertension induced neovascular sequelae, or chronic liver
XX infection. The peptides are also used to modulate vascular permeability
XX in a mammal (the mammal has a condition characterised by fluid
XX accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX or brain. The peptides are used to image blood vessels and lymphatic

```

CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

XX Sequence 17 AA;

Query Match 90.3%; Score 65; DB 22; Length 17;
 Best Local Similarity 84.6%; Pred. NO. 0.00021;
 Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASELGKSTNTEPC 13
 Db 1 case1grstnsfc 13

```

RESULT 11
AAU04536
ID AAU04536 standard; Peptide; 13 AA.
XX
AC AAU04536;
XX
DT 26-SEP-2001 (first entry)
XX
DE VEGF based monocyclic peptide 14.
XX
KW Human: VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..13 /note= "This bond cyclises the peptide"
FT Disulfide-bond 17 /note= "This bond cyclises the peptide"
XX
PN WO200152875-A1.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01533.
XX
PR 18-JAN-2000; 2000US-0176293.
PR 16-MAY-2000; 2000US-0204590.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX
DR WPI: 2001-442248/47.
XX
PT Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
PS Example 25; Page 47; 102pp; English.
XX
XX The sequence represents a monomeric monocyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
XX VEGFD (vascular endothelial growth factor). The invention relates to a
XX method of producing a monomeric monocyclic peptide by a measuring
XX beta-beta carbon separation distances on opposite antiparallel strands of
XX a peptide loop fragment from an exposed loop of a growth factor protein
XX and cyclising the peptide by oxidising the cysteine residues. The
XX monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX monocyclic peptides) and a cyclic peptide with at least one amino acid
XX deleted prior to cyclisation are used to interfere with angiogenesis,
XX neovascularisation or lymphangiogenesis in a mammal with a condition
XX characterised by angiogenesis, neovascularisation or lymphangiogenesis.

```

CC The condition is diabetic retinopathy, psoriasis, arthropathy.
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 SO Sequence 13 AA;

Query Match 84.7% Score 61; DB 22; Length 13;
 Best Local Similarity 76.9% Pred. No. 0.00075;
 Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASELGKSTNTFC 13
 |||:|||||:|
 Db 1 casdvgykstntwc 13

RESULT 12
 AAU04548
 ID AAU04548 standard; Peptide; 17 AA.
 AC AAU04548;
 XX
 DT 26-SEP-2001 (first entry)

XX VEGF based bicyclic dimeric peptide #5.
 DE
 XX

KW Human: VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..13
 FT /note- "This bond cyclises the peptide"
 FT Disulfide-bond 17
 FT /note- "A disulfide bond forms between residue
 FT 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide"
 XX

PN WO200152875-A1.

XX 26-JUL-2001.

PD 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX

PS Example 26; Page 49; 102pp; English.

XX The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF-D (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC a beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 SO Sequence 17 AA;

Query Match 84.7% Score 61; DB 22; Length 17;
 Best Local Similarity 76.9% Pred. No. 0.001;
 Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASELGKSTNTFC 13
 |||:|||||:|
 Db 1 casdvgykstntwc 13

RESULT 13
 AAU04522
 ID AAU04522 standard; protein; 73 AA.
 XX
 AC AAU04522;
 XX

DT 26-SEP-2001 (first entry)

XX Human VEGF-D amino acids Val101-Thr 173.

DE Human: VEGF-D; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation.

XX Homo sapiens.

PN WO200152875-A1.

XX 26-JUL-2001.

PD 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

XX Example 1; Page 90-91; 102pp; English.

XX The sequence represents Human VEGF-D (vascular endothelial growth factor)
CC amino acids Val101-Thr 173, used together with the C-terminal 23
CC residues of VEGF to make a hybrid theoretical molecule for 3
CC dimensional modelling. The sequence is used in a method of producing
CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
CC distances on opposite antiparallel strands of a peptide loop fragment
CC from an exposed loop of a growth factor protein and cyclising the peptide
CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
CC bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic
CC peptide with at least one amino acid deleted prior to cyclisation are
CC used to interfere with angiogenesis, neovascularisation or
CC lymphangiogenesis in a mammal with a condition characterised by
CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
CC malignant or benign tumour, post-recovery cerebrovascular accident,
CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
CC neovascularisation of the liver, excessive hormone-related angiogenic
CC dysfunction, diabetes induced neovascular sequelae, hypertension induced
CC neovascular sequelae, or chronic liver infection. The peptides are also
CC used to modulate vascular permeability in a mammal (the mammal has a
CC condition characterised by fluid accumulation in peripheral limbs or in
CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
CC peptides are used to interfere with at least one biological activity
CC induced by VEGF, VEGF-C or -D and are also used in combination with an
CC anti-inflammatory agent, to treat a chronic inflammation, especially
CC rheumatoid arthritis, psoriasis and diabetic retinopathy.

XX Sequence 73 AA;

XX Query Match 75.0%; Score 54; DB 22; Length 73;
XX Best Local Similarity 100.0%; Pred. No. 0.072;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTF 12
DB 21 aselgkstntf 31

RESULT 14

ID AAU04520 standard; protein; 96 AA.

XX AAU04520;

DT 26-SEP-2001 (first entry)

XX Human VEGF-D amino acids Val101-PRO186.

XX Human; VEGF-D; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation.

OS Homo sapiens.

XX WO200152875-A1.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01533.

PR 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204550.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI; 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

XX Example 1; Page 89; 102pp; English.

XX The sequence represents Human VEGF-D (vascular endothelial growth factor)
CC amino acids Val101-PRO186. The sequence is used in a method of producing
CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
CC distances on opposite antiparallel strands of a peptide loop fragment
CC from an exposed loop of a growth factor protein and cyclising the peptide
CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
CC bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic
CC peptide with at least one amino acid deleted prior to cyclisation are
CC used to interfere with angiogenesis, neovascularisation or
CC lymphangiogenesis in a mammal with a condition characterised by
CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
CC malignant or benign tumour, post-recovery cerebrovascular accident,
CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
CC neovascularisation of the liver, excessive hormone-related angiogenic
CC dysfunction, diabetes induced neovascular sequelae, hypertension induced
CC neovascular sequelae, or chronic liver infection. The peptides are also
CC used to modulate vascular permeability in a mammal (the mammal has a
CC condition characterised by fluid accumulation in peripheral limbs or in
CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
CC peptides are used to interfere with at least one biological activity
CC induced by VEGF, VEGF-C or -D and are also used in combination with an
CC anti-inflammatory agent, to treat a chronic inflammation, especially
CC rheumatoid arthritis, psoriasis and diabetic retinopathy.

XX Sequence 96 AA;

XX Query Match 75.0%; Score 54; DB 22; Length 96;
XX Best Local Similarity 100.0%; Pred. No. 0.096;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTF 12
DB 21 aselgkstntf 31

RESULT 15

ID AAY23889 standard; protein; 109 AA.

XX AAY23889;

DT 21-SEP-1999 (first entry)

XX Human vascular endothelial growth factor (VEGF)-D.

XX Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
KW tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
KW wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.

OS Homo sapiens.

XX WO9933485-A1.

PD 08-JUL-1999.

PF 23-DEC-1998; 98WO-US27373.

XX 29-MAY-1998; 98US-0087392.
PR 24-DEC-1997; 97AU-0001131.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Alltalo K, Stacker SA:
XX
XX WPI; 1999-405368/34.
XX
XX A human cell line stably expressing vascular endothelial growth
PT factor D, useful for treating melanomas or tumours expressing VEGF-D
XX
XX
PS Claim 6; Page 72; 79pp; English.
XX
XX The present sequence represents human vascular endothelial growth
CC factor (VEGF)-D. The specification describes a human cell line
CC which stably expresses VEGF-D, or fragments/analogues having VEGF-D
CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
CC antibodies), are useful for the treatment or alleviation of malignant
CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
XX
SQ Sequence 109 AA;

Query Match 75.0%; Score 54; DB 20; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTF 12
|||
Db 29 aselgkstntf 39

Search completed: June 17, 2002, 16:02:11
Job time: 420 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:55:11 ; Search time 75.98 Seconds
(without alignments)
4.179 Million cell updates/sec

Title: US-09-761-636A-5
Perfect score: 72
Sequence: 1 CASELCKSTNTFC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	75.0	325	4 US-08-915-795-3	Sequence 3, Appl
2	54	75.0	354	4 US-08-915-795-5	Sequence 5, Appl
3	51	70.8	321	4 US-08-915-795-9	Sequence 9, Appl
4	51	70.8	356	4 US-08-915-795-8	Sequence 8, Appl
5	43	59.7	83	3 US-08-875-811-2	Sequence 1, Appl
6	43	59.7	104	1 US-08-283-971-1	Sequence 1, Appl
7	43	59.7	104	1 US-07-921-619-1	Sequence 1, Appl
8	43	59.7	104	1 US-08-467-955-2	Sequence 1, Appl
9	43	59.7	104	1 US-08-467-955-2	Sequence 1, Appl
10	43	59.7	104	2 US-08-891-848-13	Sequence 13, Appl
11	43	59.7	104	3 US-08-875-811-1	Sequence 1, Appl
12	43	59.7	104	4 US-09-394-268-1	Sequence 1, Appl
13	43	59.7	104	4 US-09-394-268-2	Sequence 2, Appl
14	43	59.7	105	3 US-08-875-811-24	Sequence 26, Appl
15	43	59.7	105	3 US-08-875-811-26	Sequence 26, Appl
16	43	59.7	105	3 US-08-875-811-39	Sequence 39, Appl
17	43	59.7	106	3 US-08-875-811-28	Sequence 28, Appl
18	43	59.7	107	3 US-08-875-811-20	Sequence 20, Appl
19	43	59.7	107	3 US-08-875-811-30	Sequence 30, Appl
20	43	59.7	111	3 US-08-875-811-22	Sequence 22, Appl
21	43	59.7	112	3 US-08-875-811-32	Sequence 32, Appl
22	43	59.7	129	3 US-08-875-811-63	Sequence 63, Appl
23	43	59.7	251	3 US-08-875-811-59	Sequence 59, Appl
24	43	59.7	254	3 US-08-875-811-61	Sequence 61, Appl
25	43	59.7	355	3 US-08-875-811-41	Sequence 41, Appl
26	43	59.7	355	3 US-08-875-811-49	Sequence 49, Appl
27	43	59.7	355	3 US-08-875-811-57	Sequence 57, Appl

28	43	59.7	355	3 US-08-875-811-64	Sequence 64, Appl
29	43	59.7	358	3 US-08-875-811-51	Sequence 51, Appl
30	43	59.7	366	3 US-08-875-811-55	Sequence 55, Appl
31	43	59.7	379	3 US-08-875-811-43	Sequence 43, Appl
32	38	52.8	496	3 US-08-889-841B-12	Sequence 12, Appl
33	38	52.8	496	3 US-08-889-841B-16	Sequence 16, Appl
34	37	51.4	169	2 US-08-790-032-2	Sequence 2, Appl
35	37	51.4	169	2 US-09-069-619-2	Sequence 2, Appl
36	36	50.0	371	4 US-09-199-637A-295	Sequence 295, App
37	35	48.6	27	1 US-08-066-325-83	Sequence 83, Appl
38	35	48.6	45	4 US-08-857-076-62	Sequence 62, Appl
39	35	48.6	45	4 US-08-857-076-92	Sequence 92, Appl
40	35	48.6	249	2 US-08-989-386-5	Sequence 5, Appl
41	35	48.6	509	4 US-08-472-240A-20	Sequence 20, Appl
42	35	48.6	610	4 US-09-257-490-12	Sequence 12, Appl
43	35	48.6	855	4 US-07-956-483-15	Sequence 15, Appl
44	35	48.6	887	4 US-08-472-240A-6	Sequence 6, Appl
45	35	48.6	1018	5 PCT-US93-11703-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1
US-08-915-795-3
Sequence 3, Application US/08915795

Patent No. 6235713
GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER

APPLICANT: Karl ALITALO
TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Evenson, McKeeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700

CITY: Washington
STATE: DC

COUNTRY: United States of America
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795
FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844

TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: NO
ORIGINAL SOURCE:

TISSUE TYPE: Human Breast
US-08-915-795-3

Query Match 75.0%; Score 54; DB 4; Length 325;
Best local Similarity 100.0%; Pred. No. 0.079;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTE 12
Db 92 ASELGKSTNTE 102

RESULT 2

US-08-915-795-5
Sequence 5, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Lung
US-08-915-795-5

Query Match

Best Local Similarity 75.0%; Score 54; DB 4; Length 354;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTE 12
Db 121 ASELGKSTNTE 131

RESULT 3

US-08-915-795-9
Sequence 9, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

TISSUE TYPE: Mouse Lung

US-08-915-795-9

Query Match

Best Local Similarity 70.8%; Score 51; DB 4; Length 321;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTE 12
Db 121 ASELGKSTNTE 131

RESULT 4

US-08-915-795-8
Sequence 8, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Mouse Lung
US-08-915-795-8

Query Match 70.8%; Score 51; DB 4; Length 358;
Best Local Similarity 90.9%; Pred. No. 0.28;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASELGKSTNTF 12
|||||:||||
Db 126 ASELGKSTNTF 136

RESULT 5
US-08-875-811-2
Sequence 2, Application US/08675811
Patent No. 6045793
GENERAL INFORMATION:
APPLICANT: Rybak, Susanna M.
APPLICANT: Newton, Dianne L.
APPLICANT: Bogue, Luis
APPLICANT: Wlodawer, Alexander
TITLE OF INVENTION: Recombinant Ribonuclease Proteins
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,811
FILING DATE: 19-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US97/02588
FILING DATE: 19-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,800
FILING DATE: 21-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Paris, Susan K.
REGISTRATION NUMBER: 41,739
REFERENCE/DOCKET NUMBER: 015280-244100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..83
OTHER INFORMATION: /note="Rana clone 9 peptide from Rana
OTHER INFORMATION: plipens genomic DNA"
US-08-875-811-2

Query Match 59.7%; Score 43; DB 3; Length 83;
Best Local Similarity 61.5%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNTFC 13
|:|||||
Db 60 CKYKLKSTNKFC 72

RESULT 6
US-08-283-971-1
Sequence 1, Application US/08283971
Patent No. 5529775
GENERAL INFORMATION:
APPLICANT: Ardeit Ph.D, Wojciech J.
APPLICANT: Mikulski, Stanislaw M.
TITLE OF INVENTION: PHARMACEUTICAL FOR TREATING TUMORS IN HUMANS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mark H. Jay, P.C.
STREET: P.O. Box 020083, General Post Office
CITY: Brooklyn
STATE: New York
COUNTRY: USA
ZIP: 11202-0002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,971
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/921,180
FILING DATE: 30-JUL-1992
APPLICATION NUMBER: US 07/178,118
FILING DATE: 06-APR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/436,141
FILING DATE: 13-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Jay, Mark H.
REGISTRATION NUMBER: 27507
REFERENCE/DOCKET NUMBER: 5006 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 718-625-0399
TELEFAX: 718-625-0399
TELEX: No. 5529775 Applicable
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:

ORGANISM: Rana pipiens
DEVELOPMENTAL STAGE: Embryo
US-08-283-971-1

Query Match 59.7%; Score 43; DB 1; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 CASELGKSTNFC 13
Db 75 CKYKLRKSTNFC 87

RESULT 7
US-07-921-619-1
; Sequence 1, Application US/07921619
; Patent No. 5595734
; GENERAL INFORMATION:
; APPLICANT: Ardelt Ph.D, Wojciech J.
; TITLE OF INVENTION: PHARMACEUTICAL FOR TREATING TUMORS IN HUMANS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark H. Jay, P.C.
; STREET: P.O. Box 020083, General Post Office
; CITY: Brooklyn
; STATE: New York
; COUNTRY: USA
; ZIP: 11202-0002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921,619
; FILING DATE: 19920728
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/178,118
; FILING DATE: 06-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/436,141
; FILING DATE: 13-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Jay, Mark H.
; REGISTRATION NUMBER: 27507
; REFERENCE/DOCKET NUMBER: 5005 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 718-625-0399
; TELEFAX: 718-625-0399
; TELEX: NO. 5595734 Applicable
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rana pipiens
; DEVELOPMENTAL STAGE: Embryo
US-07-921-619-1

Query Match 59.7%; Score 43; DB 1; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNFC 13
Db 75 CKYKLRKSTNFC 87

RESULT 8
US-08-467-955-1
; Sequence 1, Application US/08467955
; Patent No. 572805
; GENERAL INFORMATION:
; APPLICANT: Ardelt Ph.D, Wojciech J.
; TITLE OF INVENTION: PHARMACEUTICALS AND METHOD FOR MAKING THEM
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark H. Jay, P.A.
; STREET: P.O. Box E
; CITY: Short Hills
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07078-0383
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,955
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/178,118
; FILING DATE: 06-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/436,141
; FILING DATE: 13-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/814,332
; FILING DATE: 03-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,970
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jay, Mark H.
; REGISTRATION NUMBER: 27507
; REFERENCE/DOCKET NUMBER: 5007 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-912-9066
; TELEFAX: 201-912-0442
; TELEX: NO. 5728805 Applicable
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rana pipiens
; DEVELOPMENTAL STAGE: Oocyte
US-08-467-955-1

Query Match 59.7%; Score 43; DB 1; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 CASELGKSTNFC 13
Db 75 CKYKLRKSTNFC 87

RESULT 9
US-08-467-955-2
; Sequence 2, Application US/08467955
; Patent No. 5728805
; GENERAL INFORMATION:
; APPLICANT: Argelt Ph.D, Wojciech J.
; TITLE OF INVENTION: PHARMACEUTICALS AND METHOD FOR MAKING THEM
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark H. Jay, P.A.
; STREET: P.O. Box E
; CITY: Short Hills
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07078-0383
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,955
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/178,118
; FILING DATE: 06-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/436,141
; FILING DATE: 13-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/814,332
; FILING DATE: 03-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,970
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jay, Mark H.
; REGISTRATION NUMBER: 27507
; REFERENCE/DOCKET NUMBER: 5007 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-912-9066
; TELEFAX: 201-912-0442
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Rana pipiens
; DEVELOPMENTAL STAGE: Oocyte
; US-08-467-955-2

Query Match 59.7%; Score 43; DB 1; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFC 13
I : | | | | |
Db 75 CKYIKKSTNKFC 87

RESULT 10
US-08-891-848-13
; Sequence 13, Application US/08891848

; Patent No. 5955073
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Youle, Richard J.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Nicholls, Peter J.
; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,848
; FILING DATE: NO. 5955073 yet assigned
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,462
; FILING DATE: 22-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/014,082
; FILING DATE: 04-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,195
; FILING DATE: 22-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/510,696
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-110310US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..104
; OTHER INFORMATION:
; OTHER INFORMATION: /note="Onconase from Rana pipiens"
; US-08-891-848-13

Query Match 59.7%; Score 43; DB 2; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFC 13
I : | | | | |
Db 75 CKYIKKSTNKFC 87

RESULT 11
US-08-875-811-1
; Sequence 1, Application US/08875811
; Patent No. 6045793
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.

APPLICANT: Newton, Dianne L.
APPLICANT: Boque, Luis
TITLE OF INVENTION: Recombinant Ribonuclease Proteins
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,811
FILING DATE: 19-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/02588
FILING DATE: 19-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,800
FILING DATE: 21-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Paris, Susan K.
REGISTRATION NUMBER: 41,739
REFERENCE/DOCKET NUMBER: 015280-244100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..104
OTHER INFORMATION: /label=nonc
OTHER INFORMATION: /note="native ONCOMASE (Registered
OTHER INFORMATION: Trademark) from Rana pipiens"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="Xaa - pyroglutamic acid"
US-08-875-811-1

Query Match 59.7%; Score 43; DB 3; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELCKSTNFC 13
DB 75 CKYKLNKSTNFC 87

RESULT 12
US-09-394-268-1
Sequence 1, Application US/09394268
Patent No. 6175003
GENERAL INFORMATION:
APPLICANT: Saxena, Shalendra K
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING RIBONUCLEASES AND METHODS OF
FILE REFERENCE: MAKING THEM
CURRENT APPLICATION NUMBER: US/09/394,268

CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 104
TYPE: PRT
ORGANISM: Rana pipiens
US-09-394-268-1

Query Match 59.7%; Score 43; DB 4; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELCKSTNFC 13
DB 75 CKYKLNKSTNFC 87

RESULT 13
US-09-394-268-2
Sequence 2, Application US/09394268
Patent No. 6175003
GENERAL INFORMATION:
APPLICANT: Saxena, Shalendra K
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING RIBONUCLEASES AND METHODS OF
FILE REFERENCE: 5013
CURRENT APPLICATION NUMBER: US/09/394,268
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 104
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: SEQ ID NO:1 with Leu at position 23 and Cys at
OTHER INFORMATION: position 72
US-09-394-268-2

Query Match 59.7%; Score 43; DB 4; Length 104;
Best Local Similarity 61.5%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELCKSTNFC 13
DB 75 CKYKLNKSTNFC 87

RESULT 14
US-08-875-811-24
Sequence 24, Application US/08875811
Patent No. 6045793
GENERAL INFORMATION:
APPLICANT: Rybak, Susanna M.
APPLICANT: Boque, Luis
APPLICANT: Newton, Dianne L.
APPLICANT: Wlodawer, Alexander
TITLE OF INVENTION: Recombinant Ribonuclease Proteins
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/875,811
 FILING DATE: 19-FEB-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US97/02588
 FILING DATE: 19-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/011,800
 FILING DATE: 21-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Faris, Susan K.
 REGISTRATION NUMBER: 41,739
 REFERENCE/DOCKET NUMBER: 015280-244100US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 105 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-875-811-24

Query Match 59.7%; Score 43; DB 3; Length 105;
 Best Local Similarity 61.5%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNFC 13
 | : | | | | | |
 Db 76 CKYKLKSTNFC 88

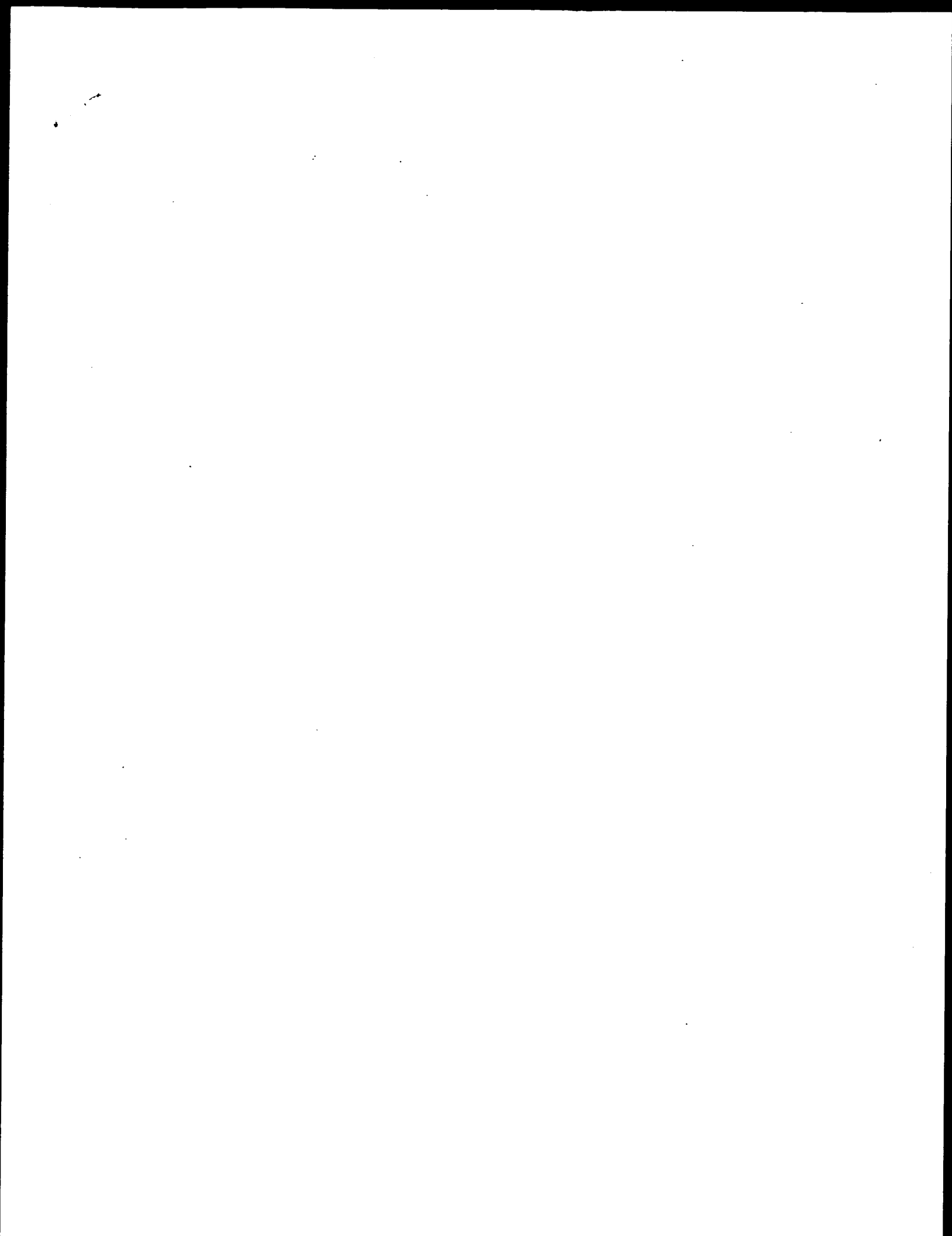
RESULT 15
 US-08-875-811-26
 Sequence 26, Application US/08875811
 Patent No. 6045793
 GENERAL INFORMATION:
 APPLICANT: Rybak, Susanna M.
 APPLICANT: Newton, Dianne L.
 APPLICANT: Boque, Luis
 APPLICANT: Wlodawer, Alexander
 TITLE OF INVENTION: Recombinant Ribonuclease Proteins
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/875,811
 FILING DATE: 19-FEB-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US97/02588
 FILING DATE: 19-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/011,800
 FILING DATE: 21-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Faris, Susan K.
 REGISTRATION NUMBER: 41,739
 REFERENCE/DOCKET NUMBER: 015280-244100US

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 105 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-875-811-26

Query Match 59.7%; Score 43; DB 3; Length 105;
 Best Local Similarity 61.5%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CASELGKSTNFC 13
 | : | | | | | |
 Db 76 CKYKLKSTNFC 88

Search completed: June 17, 2002, 15:56:35
 Job time: 84 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:18 ; Search time 95.43 Seconds
(without alignments)
8.055 Million cell updates/sec

Title: US-09-761-636A-6
Perfect score: 46
Sequence: 1 CNEESLIC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	78.3	120	2	A33787
2	36	78.3	133	2	B49530
3	36	78.3	146	2	S57956
4	36	78.3	190	2	B40080
5	36	78.3	1115	2	A45761
6	35	76.1	4427	2	PN0637
7	34.5	75.0	351	2	T42421
8	34.5	75.0	1778	2	T50074
9	34	73.9	295	2	H71078
10	34	73.9	501	1	A40938
11	34	73.9	642	2	T51421
12	34	73.9	1722	1	I78879
13	33	71.7	190	2	B44881
14	33	71.7	190	2	A35987
15	33	71.7	214	2	A44881
16	33	71.7	291	2	AF0123
17	33	71.7	325	2	G90130
18	33	71.7	350	2	D69130
19	33	71.7	551	2	JC7562
20	33	71.7	1786	1	MMH081
21	32	69.6	169	2	S76289
22	32	69.6	183	2	E72459
23	32	69.6	190	2	S52130
24	32	69.6	232	2	A41551
25	32	69.6	272	2	S13587
26	32	69.6	290	2	G97328
27	32	69.6	338	2	E86465
28	32	69.6	358	2	A82206
29	32	69.6	376	2	B72321

30	32	69.6	399	2	H87486	1-deoxy-D-xylose
31	32	69.6	429	2	JH0151	malate dehydrog
32	32	69.6	429	2	S13586	malate dehydrog
33	32	69.6	429	2	T16656	hypothetical prote
34	32	69.6	430	2	S17781	malate dehydrog
35	32	69.6	432	1	DEM2MC	malate dehydrog
36	32	69.6	432	2	S20743	malate dehydrog
37	32	69.6	435	2	S52268	malate dehydrog
38	32	69.6	441	2	S33066	hypothetical prote
39	32	69.6	441	2	S38346	hypothetical prote
40	32	69.6	463	2	T15416	probable kinase yp
41	32	69.6	765	2	AF0118	hypothetical prote
42	32	69.6	766	2	AG0179	glycosyltransfer
43	32	69.6	1044	2	H97186	hypothetical prote
44	32	69.6	1047	2	T21306	protein F23B2.4 [1
45	32	69.6	1124	2	H88772	

ALIGNMENTS

RESULT 1
A33787
vascular endothelial growth factor (version 1) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 05-Nov-1999
C:Accession: A33787
R:Rischer, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.
Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989
A:Title: Vascular endothelial growth factor: a new member of the platelet-derived
A:Reference number: A33787; MUID:90121225
A:Accession: A33787
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <TIS>
A:Cross-references: GB:M3750; NID:q163810; PIDN:AAA30805.1; PID:q163811
C:Keywords: alternative splicing

Query Match 78.3%; Score 36; DB 2; Length 120;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
DB 60 CNEESLIC 67

RESULT 2
B49530
vascular endothelial growth factor homolog A2R, 14.7K - Orf virus

C:Species: Orf virus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: B49530
R:Lytle, D.J.; Fraser, K.M.; Fleming, S.B.; Mercer, A.A.; Robinson, A.J.
J. Virol. 68, 84-92, 1994
A:Title: Homologs of vascular endothelial growth factor are encoded by the poxv
A:Reference number: A49530; MUID:94076465
A:Contents: N22
A:Accession: B49530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <LTY>
A:Cross-references: GB:S67520; NID:q456897; PIDN:AAB29220.1; PID:q456899
A:Note: sequence inconsistent with nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:141420, NCBIR:141425)

Query Match 78.3%; Score 36; DB 2; Length 133;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8

Db 71 CNDESLC 78

RESULT 3

S57956
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C/Accession: S57956
R:Reimer, D.A.; Dai, Y.; Li, J.; Jones, S.C.; Moor, R.M.
submitted to the EMBL Data Library, July 1995
A:Reference number: S57956
A:Accession: S57956
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RED>
A:Cross-references: EMBL:X89506; NID:g899350; PIDN:CAA61677.1; PID:g899351

Query Match
Best Local Similarity 78.3%; Score 36; DB 2; Length 146;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESLITC 8
Db 86 CNDESLC 93

RESULT 4

B40080
C:Species: Bos primigenius tauros (cattle)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Nov-1999
C/Accession: B40080; B33787; A33255
R:Leung, D.W.; Cachianes, G.; Kuang, W.J.; Goeddel, D.V.; Ferrara, N.
Science 246, 1306-1309, 1989
A>Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.
A:Reference number: A40080; MUID:90069608
A:Accession: B40080
A:Molecule type: mRNA
A:Residues: 1-190 <LEU>
A:Cross-references: GB:M32976; NID:q163006; PIDN:AAA30502.1; PID:q163007
R:Tischer, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.; Crist
Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989
A>Title: Vascular endothelial growth factor: a new member of the platelet-derived growth
A:Reference number: A33787; MUID:90121225
A:Accession: B33787
A:Molecule type: mRNA
A:Residues: 27-190 <TR>
A:Cross-references: GB:M31836; NID:q163808; PIDN:AAA30804.1; PID:q163809
R:Ferrara, N.; Henzel, W.J.
Biochem. Biophys. Res. Commun. 161, 851-858, 1989
A>Title: Pituitary follicular cells secrete a novel heparin-binding growth factor specific
A:Reference number: A33255; MUID:89286596
A:Accession: A33255
A:Molecule type: protein
A:Residues: 27-31 <FPR>
C:Keywords: alternative splicing; glycoprotein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-190/Product: vascular endothelial growth factor #status predicted <MAT>
F:100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match
Best Local Similarity 78.3%; Score 36; DB 2; Length 190;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESLITC 8
Db 86 CNDESLC 93

RESULT 5

A45761
Ca2+-transporting ATPase (EC 3.6.1.38) - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 22-Jun-1999
C/Accession: A45761
R:Murakami, K.; Tanabe, K.; Takada, S.
J. Cell Sci. 97, 487-495, 1990
A>Title: Structure of a Plasmodium yoelii gene-encoded protein homologous to the
A:Reference number: A45761; MUID:91161669
A:Accession: A45761
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1115 <MUR>
A:Cross-references: GB:X55197; NID:q10097; PIDN:CAA38962.1; PID:q10098
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding
C:Keywords: hydrolase
F:689-866/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match
Best Local Similarity 78.3%; Score 36; DB 2; Length 1115;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESLITC 8
Db 486 CNESLITC 493

RESULT 6

PN0637
polyketide synthase pksL - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 19-May-1994 #sequence_revision 06-Feb-1995 #text_change 03-Nov-2000
C/Accession: S25021; PN0637; B69679
R:Scotti, C.; Platt, M.; Cuzon, A.; Tognoni, A.; Grandi, G.; Gallizzi, A.; Alber
submitted to the EMBL Data Library, July 1992
A:Description: A Bacillus subtilis large ORF coding for a polypeptide highly simil
A:Reference number: S25021
A:Accession: S25021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4427 <SCOT>
A:Cross-references: EMBL:Z14098; NID:q40057; PIDN:CAA78479.1; PID:q40058
R:Scotti, C.; Platt, M.; Cuzon, A.; Perant, P.; Tognoni, A.; Grandi, G.; Gallizzi
Gene 130, 65-71, 1993
A>Title: A Bacillus subtilis large ORF coding for a polypeptide highly similar to
A:Reference number: PN0637; MUID:93345824
A:Accession: PN0637
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 184-282;382-850;926-1115;1409-1648;1665-1761;1876-2344;2469-2560;2609-
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; I
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fuma, S.; Gallizzi, A.; G
Iech, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardi
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; M
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Port
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, J.; Sato, T.; Sca
A:Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.;
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosto, V.; Uchi
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshi
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033
A:Accession: B69679
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-4427 <KUN>
A:Cross-references: GB:Z99113; GB:AL009126; NID:q2634090; PIDN:CAI3602.1; PID:q263

A:Experimental source: strain 168
 C:Comment: This enzyme is composed of four synthase units. Unit1 comprises beta-ketosynth acyl-carrier protein domains. Unit3 comprises beta-ketosynthase, acyl-carrier protein and C:Genetics:
 A:Gene: pksL
 C:Superfamily: Bacillus subtilis polyketide synthase pksL; 3-oxoacyl-[acyl-carrier-prote
 C:Keywords: acyltransferase; carrier protein
 F:343-758/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:1410-1591/Domain: short-chain alcohol dehydrogenase homology <SAD1>
 F:1836-2252/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:2485-2559/Domain: acyl carrier protein homology <ACPI>
 F:2626-2700/Domain: acyl carrier protein homology <ACPI>
 F:2783-3181/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F:3576-3774/Domain: short-chain alcohol dehydrogenase homology <SAD2>
 F:3852-3922/Domain: acyl carrier protein homology <ACP3>
 F:3992-4372/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>

Query Match 76.1%; Score 35; DB 2; Length 4427;
 Best Local Similarity 62.5%; Pred. No. 3; 7e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 |||||
 Db 1205 CNEBQVC 1212

RESULT 7
 T42421
 hypothetical protein - fission yeast (Schizosaccharomyces pombe) (fragment)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T42421
 R:Yoshida, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: Z17323; MUID:98162722
 A:Accession: T42421
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-351 <YOS>
 A:Cross-references: EMBL:DB9145; NID:91749497; PIDN:BAA13807.1; PID:91749498
 A:Experimental source: strain PR745

Query Match 75.0%; Score 34.5; DB 2; Length 351;
 Best Local Similarity 72.7%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 1 CNEESL---IC 8
 |||||
 Db 321 CNEESLTDATC 331

RESULT 8
 T50074
 probable nucleoporin [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
 C:Accession: T50074
 R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25034
 A:Accession: T50074
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1778 <MCD>
 A:Cross-references: EMBL:AL133357; PIDN:CAB62415.1; GSPDB:GN00066; SPDB:SPAC1486.05
 C:Genetics:
 A:Gene: SPDB:SPAC1486.05
 A:Map position: 1
 A:Introns: 139/2

Query Match 75.0%; Score 34.5; DB 2; Length 1778;
 Best Local Similarity 72.7%; Pred. No. 2; 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 1 CNEESL---IC 8
 |||||
 Db 1748 CNEESLTDATC 1758

RESULT 9
 H71078
 probable cation efflux system protein czcd - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
 C:Accession: H71078
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.;
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermo
 A:Reference number: A71000; MUID:98344137
 A:Accession: H71078
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-295 <KAW>
 A:Cross-references: GB:AP000004; NID:93236131; PIDN:BAA29990.1; PID:93257307
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by G
 C:Genetics:
 A:Gene: PH0896
 C:Superfamily: zinc transporter Znt-2

Query Match 73.9%; Score 34; DB 2; Length 295;
 Best Local Similarity 62.5%; Pred. No. 56;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 |||||
 Db 283 CKEKSMIC 290

RESULT 10
 A40938
 cytochrome P450 1b - rabbit
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: A40938
 R:Kikuta, Y.; Sogawa, K.; Hanlu, M.; Kinoshita, M.; Kusunose, E.; Nojima, Y.; Yam
 J. Biol. Chem. 266, 17821-17825, 1991
 A:Title: A novel species of cytochrome P-450 (P-450-1b) specific for the small
 A:Reference number: A40938; MUID:92011499
 A:Accession: A40938
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-501 <KIK>
 A:Cross-references: GB:D90405; NID:9217717; PIDN:BAA14401.1; PID:9217718
 C:Genetics:
 A:Gene: CYP2J1
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:308-469/Domain: cytochrome P450 homology <CYP>
 F:447/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 73.9%; Score 34; DB 1; Length 501;
 Best Local Similarity 85.7%; Pred. No. 89;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 NEESLIC 8
 |||||

Db 297 NEENLIC 303

RESULT 11

T51421

L:aspartate oxidase-like protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000

C:Accession: T51421

R:Salto, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

A:Reference number: 225394

A:Accession: T51421

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <SAT>

A:Cross-references: EMBL:AL391149

A:Experimental source: cultivar Columbia; BAC clone T9L3

C:Genetics:

A:Map position: 5

A:Introns: 43/2; 158/3; 239/3; 258/3; 278/3; 299/2

A:Note: T9L3_60

C:Superfamily: fumarate reductase flavoprotein; 3-oxosteroid 1-dehydrogenase homology; F

Query Match

Best Local Similarity 73.9%; Score 34; DB 2; Length 642;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNEESLIC 8

1:111111

Db 154 CDEETVVC 161

RESULT 12

T78879

retinoblastoma binding protein 2 - human

C:Species: Homo sapiens (man)

C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: T78879; S16954

R:Faithy, A.R.; Helin, K.; Dembski, M.S.; Dyson, N.; Harlow, E.; Vuocolo, G.A.; Hanolik

Oncogene 8; 3149-3156, 1993

A:Title: Characterization of the retinoblastoma binding proteins RBP1 and RBP2.

A:Reference number: 158383; MUID:94020841

A:Accession: T78879

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1722 <FAT>

A:Cross-references: GB:S66431; NID:9435777; PIDN:AAB28544.1; PID:9435778

R:Deleo-Jones, D.; Huang, P.S.; Jones, R.E.; Haskell, K.M.; Vuocolo, G.A.; Hanolik, M.G.

Nature 352; 251-254, 1991

A:Title: Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene p107

A:Reference number: S16953; MUID:91312450

A:Accession: S16954

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1102-1562; 'KKK' <DEF>

C:Genetics:

A:Gene: GDB:RBP2

A:Cross-references: GDB:119548; OMIM:180280

A:Map position: 3q21-3pter

C:Superfamily: human retinoblastoma binding protein 2

Query Match

Best Local Similarity 73.9%; Score 34; DB 1; Length 1722;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CNEESLIC 8

11111111

Db 692 CNBRLVC 699

RESULT 13

B44881

vascular endothelial growth factor-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 05-Nov-1999

C:Accession: B44881; A43351; A61029

R:Breier, G.; Aldred, U.; Sterrer, S.; Risau, W.

Development 114; 521-532, 1992

A:Title: Expression of vascular endothelial growth factor during embryonic angiogenesis

A:Reference number: A44881; MUID:92274860

A:Accession: B44881

A:Molecule type: mRNA

A:Residues: 1-190 <BRE>

A:Cross-references: GB:S38083; NID:9249858; PIDN:AAB22253.1; PID:9249859

A:Experimental source: embryo

A:Note: sequence extracted from NCBI backbone (NCBIN:107622, NCBIIP:107623)

R:Claffey, K.P.; Wilkison, W.O.; Spiegelman, B.M.

J. Biol. Chem. 267; 16317-16322, 1992

A:Title: Vascular endothelial growth factor. Regulation by cell differentiation and

A:Reference number: A43351; MUID:92355593

A:Accession: A43351

A:Molecule type: mRNA

A:Residues: 1-116; 'ER', 119-190 <CLA>

A:Cross-references: GB:M95200; NID:9202350; PIDN:AAA0547.1; PID:9202351

R:Rosenthal, R.A.; Megyesi, J.F.; Henzel, W.J.; Ferrara, N.; Folkman, J.

Growth Factors 4; 53-59, 1990

A:Title: Conditioned medium from mouse sarcoma 180 cells contains vascular endothelial

A:Reference number: A61029; MUID:91197543

A:Accession: A61029

A:Molecule type: protein

A:Residues: 27-38 <ROS>

C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; glycoprotein

Query Match

Best Local Similarity 71.7%; Score 33; DB 2; Length 190;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CNEESLIC 8

11111111

Db 86 CNDEALEC 93

RESULT 14

A35987

glioma-derived vascular endothelial cell growth factor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 05-Nov-1999

C:Accession: A35987

R:Conn, G.; Bayne, M.L.; Soderman, D.D.; Kwok, P.W.; Sullivan, K.A.; Palisi, T.M.;

Proc. Natl. Acad. Sci. U.S.A. 87; 2628-2632, 1990

A:Title: Amino acid and cDNA sequences of a vascular endothelial cell mitogen that

A:Reference number: A35987; MUID:90207249

A:Accession: A35987

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-190 <CON>

A:Cross-references: GB:M32167; NID:9204287; PIDN:AAA41211.1; PID:9204288

Query Match

Best Local Similarity 71.7%; Score 33; DB 2; Length 190;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CNEESLIC 8

11111111

Db 86 CNDEALEC 93

RESULT 15

A44881

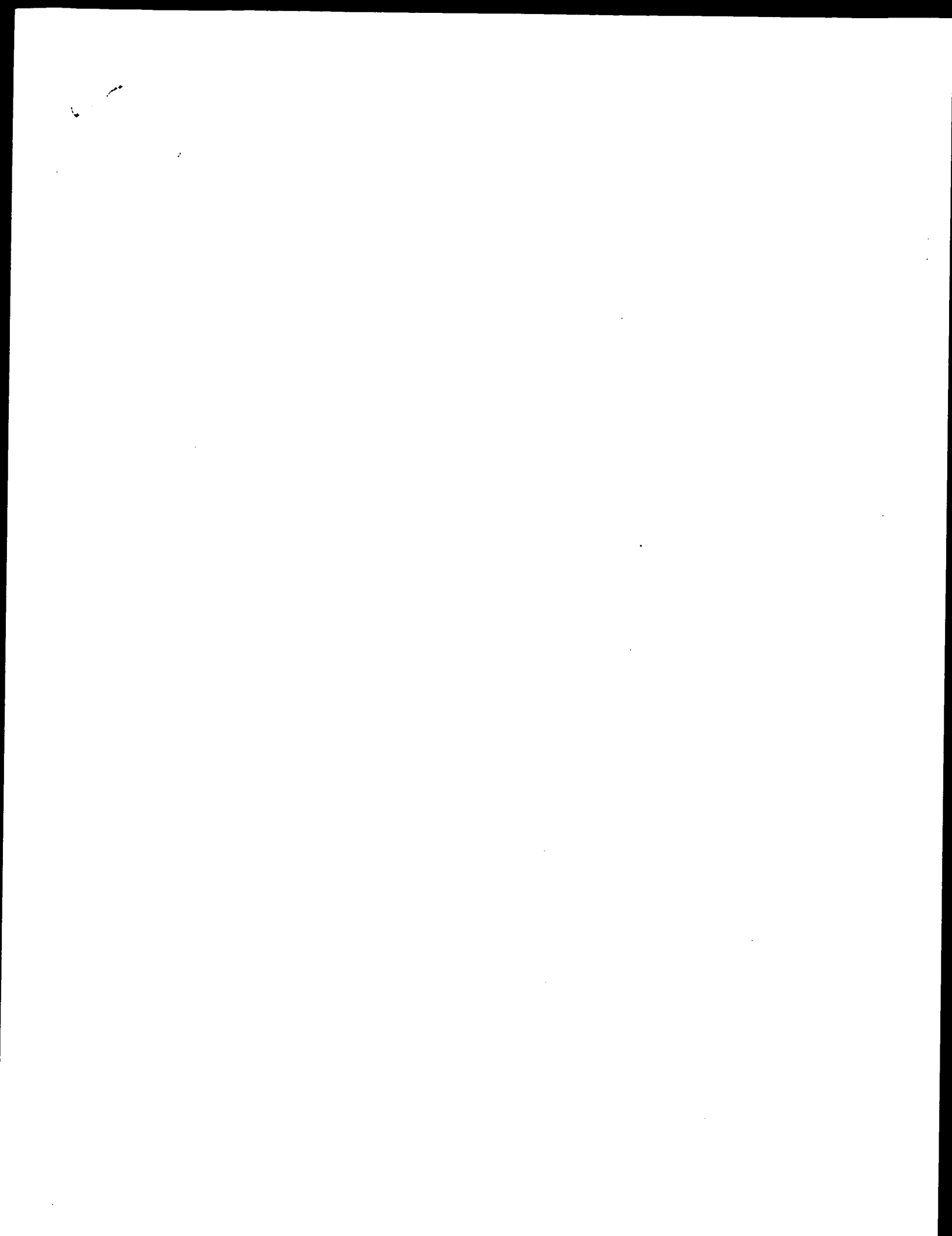
vascular endothelial growth factor-3 precursor - mouse
 N:Contains: vascular endothelial growth factor-2; vascular permeability factor
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text_change 08-Oct-1999
 C:Accession: A44881; C44881; A60932; S52136
 R:Breier, G.; Albrecht, U.; Sterrer, S.; Risau, W.
 Development 114, 521-532, 1992
 A:Title: Expression of vascular endothelial growth factor during embryonic angiogenesis
 A:Reference number: A44881; MUID:92274860
 A:Accession: A44881
 A:Molecule type: mRNA
 A:Residues: 1-214

 A:Cross-references: GB:S37052; NID:g249856; PIDN:AAB22252.1; PID:g249857
 A:Experimental source: embryo
 A:Note: sequence extracted from NCBI backbone (NCBIN:104677, NCBIP:104678)
 A:Accession: C44881
 A:Molecule type: mRNA
 A:Residues: 1-140,209-214 <BR2>
 A:Cross-references: GB:S38100; NID:g249860; PIDN:AAB22254.1; PID:g249861
 A:Note: sequence extracted from NCBI backbone (NCBIN:107624, NCBIP:107625)
 R:Clausen, M.; Gerlach, M.; Bretl, J.; Wang, F.; Familletti, P.C.; Pan, Y.C.
 J. Exp. Med. 172, 1535-1545, 1990
 A:Title: Vascular permeability factor: a tumor-derived polypeptide that induces endothelial
 A:Reference number: A60932; MUID:91079755
 A:Accession: A60932
 A:Molecule type: protein
 A:Residues: 27-33 <CLA>
 R:Sugihara, T.; Kaul, S.C.; Mitsui, Y.; Wadhwa, R.
 Biochim. Biophys. Acta 1224, 365-370, 1994
 A:Title: Enhanced expression of multiple forms of VEGF is associated with spontaneous im
 A:Reference number: S52136; MUID:95101726
 A:Accession: S52136
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 27-46 <SUG>
 C:Comment: Homodimers could be demonstrated for recombinant VEGF-2 but not VEGF-3.
 C:Keywords: alternative splicing; angiogenesis; disulfide bond; glycoprotein; homodimer;
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-214/Product: vascular endothelial growth factor-3 #status experimental <MAT>

Query Match 71.7%; Score 33; DB 2; Length 214;
 Best Local Similarity 62.5%; Pred. No. 66;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 11:1:1
 DB 86 CNDEALEC 93

Search completed: June 17, 2002, 15:58:20
 Job time: 189 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:34 ; Search time 172.85 Seconds
(without alignments)
8.007 Million cell updates/sec

Title: US-09-761-636A-6
Perfect score: 46
Sequence: 1 CNEESLIC 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL.19:*
2: SP_Archaea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_Virus:*
16: SP_Bacteriaph:*
17: SP_Archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match length	ID	Description
1	40	87.0	326 11 091ZE4	091ZE4 ratius norv
2	36	78.3	68 6 097500	097500 oryctolagus
3	36	78.3	75 6 018843	018843 oryctolagus
4	36	78.3	78 6 09N1S2	09N1S2 capreolus c
5	36	78.3	118 6 09WZB1	09WZB1 ovis aries
6	36	78.3	123 6 09N1S1	09N1S1 capreolus c
7	36	78.3	132 12 09YMF3	09YMF3 orf virus
8	36	78.3	148 13 042571	042571 xenopus lae
9	36	78.3	190 6 077643	077643 ovis aries
10	36	78.3	194 13 042572	042572 xenopus lae
11	36	78.3	421 5 095T10	095T10 drosophila
12	36	78.3	502 5 09VEF4	09VEF4 drosophila
13	36	78.3	1115 5 027764	027764 plasmodium
14	36	78.3	1341 4 096J14	096J14 homo sapien
15	34.5	75.0	351 3 P78796	P78796 schizosacch
16	34.5	75.0	1778 3 090TK4	090TK4 schizosacch

17	34	73.9	241 2 093F07	093F07 cowdria rum
18	34	73.9	295 17 058637	058637 pyrococcus
19	34	73.9	317 5 09W1E4	09W1E4 drosophila
20	34	73.9	374 11 091W71	091W71 mus musculu
21	34	73.9	435 5 09VSV3	09VSV3 drosophila
22	34	73.9	501 11 09QXF7	09QXF7 ratius norv
23	34	73.9	502 11 0924D1	0924D1 mus musculu
24	34	73.9	529 11 09CR09	09CR09 mus musculu
25	34	73.9	542 4 096PX2	096PX2 homo sapien
26	34	73.9	642 4 09LER1	09LER1 arabidopsis
27	34	73.9	838 5 09VQD9	09VQD9 drosophila
28	34	73.9	858 11 088911	088911 ratius norv
29	33	71.7	126 3 005458	005458 saccharomyc
30	33	71.7	141 11 070123	070123 mus musculu
31	33	71.7	144 13 073822	073822 brachydanio
32	33	71.7	163 10 094GR5	094GR5 oryza sativ
33	33	71.7	188 13 073682	073682 brachydanio
34	33	71.7	190 11 09QX39	09QX39 spalax leuc
35	33	71.7	190 11 091ZE1	091ZE1 ratius norv
36	33	71.7	255 11 09D7S0	09D7S0 mus musculu
37	33	71.7	325 8 098S45	098S45 guillardia
38	33	71.7	346 5 P90693	P90693 bombix mori
39	33	71.7	350 17 026438	026438 methanother
40	33	71.7	435 11 088869	088869 ratius norv
41	33	71.7	540 5 077048	077048 plasmodium
42	33	71.7	1354 5 09VKA7	09VKA7 drosophila
43	33	71.7	2479 11 063002	063002 ratius norv
44	32.5	70.7	498 10 09MAD6	09MAD6 arabidopsis
45	32.5	70.7	1015 10 09SRH8	09SRH8 arabidopsis

ALIGNMENTS

RESULT 1
ID 091ZE4 PRELIMINARY: PRT: 326 AA.
AC 091ZE4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VEGF-D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Kirkin V., Mazitschek R., Krishnan J., Steffen A., Waltenberger J.,
RA Pepper M.S., Gianni A., Sleeman J.P.;
RT "Characterization of indolones which specifically inhibit VEGF-C and
RT VEGF-D-induced activation of VEGFR-3 but not VEGFR-2.";
RL Eur. J. Biochem. 0:0-0(2001).
DR EMBL: AY032728; AAK96008.1;
SQ SEQUENCE 326 AA; 37106 MW; D7CAEBAC9FAB7D CRC64;

Query Match 87.0%; Score 40; DB 11; Length 326;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
DB 151 CNEESVWC 158
RESULT 2
ID 097500 PRELIMINARY: PRT: 68 AA.
AC 097500;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Inoue K., Kawabe Y., Kodama T.;
RT "Rabbit VEGF cDNA, partial."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR HSSP: AB020216; BAA36949.1; -
DR HSSP: P15692; 2VPF.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 68 AA; 7819 MW; 687638661E98DE0 CRC64;

Query Match
Best Local Similarity 78.3%; Score 36; DB 6; Length 68;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESSLIC 8
DB 41 CNDSESLIC 48

RESULT 3
ID 018843 PRELIMINARY; PRT; 75 AA.
AC 018843;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NEW ZEALAND WHITE; TISSUE=SKLEETAL MUSCLE;
RX Skordjanc D., Jaschinski F., Heine G., Pette D.;
RT "Sequential increases in capillarization and mitochondrial enzymes in
low-frequency-stimulated rabbit muscle."
RL Am. J. Physiol. 274:C810-C818(1998).
DR EMBL: AF022179; AAC15469.1; -
DR HSSP: P15692; 2VPF.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 75 AA; 8720 MW; DDCE2C5B29E69359 CRC64;

Query Match 78.3%; Score 36; DB 6; Length 75;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESSLIC 8

DB 29 CNDSESLIC 36
RESULT 4
ID 09M1S2 PRELIMINARY; PRT; 78 AA.
AC 09M1S2;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR ISOFORM 121 (FRAGMENT).
GN VEGF.
OS Capreolus capreolus (Roe deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Odocoileinae; Capreolus.
OX NCBI_TaxID=9858;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=20532861; PubMed=11078967;
RA Wagener A., Biotner S., Goritz F., Fickel J.;
RT "Detection of growth factors in the testis of roe deer (Capreolus
capreolus)."
RL Anim. Reprod. Sci. 64:65-75(2000).
DR EMBL: AF152393; AAF73232.1; -
DR HSSP: P15692; 2VPF.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 78 AA; 9131 MW; 7EE20DDFC17847C CRC64;

Query Match 78.3%; Score 36; DB 6; Length 78;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNESSLIC 8
DB 25 CNDSESLIC 32

RESULT 5
ID 09M2B1 PRELIMINARY; PRT; 118 AA.
AC 09M2B1;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTAL ARTERY ENDOTHELIUM;
RA Zheng J., Tsai S.C., Magness R.R.;
RT "Growth factor expression in ovine fetal placental artery endothelial
cells."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF250375; AAF75258.1; -
DR HSSP: P15692; 1VPF.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.

DR Procam; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 FT NON_TER 1
 SQ SEQUENCE 118 AA; 13931 MW; 757DC53AA56378A6 CRC64;

Query Match 78.3%; Score 36; DB 6; Length 118;
 Best Local Similarity 75.0%; Pred. No. 3.3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 11:1111
 DB 58 CNDESLIC 65

RESULT 6
 O9N1S1 PRELIMINARY; PRT; 123 AA.
 AC O9N1S1:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR ISOFORM 165 (FRAGMENT).
 GN VEGF.
 OS Capreolus capreolus (Roe deer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
 OC Cervidae; Odocoileinae; Capreolus.
 OX NCBI_TaxID=9858;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RX MEDLINE=20532861; PubMed=11078967;
 RA Wagener A., Blotner S., Goritz F., Fickel J.;
 RT "Detection of growth factors in the testis of roe deer (Capreolus capreolus).";
 RL Anim. Reprod. Sci. 64:65-75(2000).
 DR EMBL; AF152594; AAF73233.1; -.
 DR HSSP; P15692; IVP9.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 FT NON_TER 1
 FT NON_TER 123
 SQ SEQUENCE 123 AA; 14354 MW; 0A756F5A105A4CE1 CRC64;

Query Match 78.3%; Score 36; DB 6; Length 123;
 Best Local Similarity 75.0%; Pred. No. 3.4;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 11:1111
 DB 25 CNDESLIC 32

RESULT 7
 O9YMF3 PRELIMINARY; PRT; 132 AA.
 AC O9YMF3:
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR HOMOLOG VEGF-E.
 OS of virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Parapoxvirus.
 OX NCBI_TaxID=10258;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DI701;
 RX MEDLINE=99107753; PubMed=9889193;
 RA Meyer M., Claus M., Lepple-Wienhues A., Waltenberger J.,
 RA Augustin H.G., Ziche M., Lanz C., Bueltnier M., Rzila H.J., Dehio C.;
 RT "A novel vascular endothelial growth factor encoded by orf virus",
 RT VEGF-E, mediates angiogenesis via signalling through VEGFR-2 (KDR) but
 RT not VEGFR-1 (Flt-1) receptor tyrosine kinases.";
 RL EMBL J. 18:363-374(1999).
 DR EMBL; AF106020; AAD03735.1; -.
 DR HSSP; P15692; IVP9.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 SQ SEQUENCE 132 AA; 14763 MW; 15F403A068B72926 CRC64;

Query Match 78.3%; Score 36; DB 12; Length 132;
 Best Local Similarity 75.0%; Pred. No. 3.7;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 11:1111
 DB 70 CNDESLIC 77

RESULT 8
 O42571 PRELIMINARY; PRT; 148 AA.
 AC O42571:
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR 122.
 GN VEGF.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cleaver O., Tonissen K.F., Saha M.S., Krieg P.A.;
 RT "Neovascularization of the Xenopus embryo.";
 RL Dev. Dyn. 0:0-0(1997).
 DR EMBL; AF008593; AAB63679.1; -.
 DR HSSP; P15692; IVP9.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS50278; PDGF_2; 1.
 SQ SEQUENCE 148 AA; 17234 MW; 4AD153CA2F081E95 CRC64;

Query Match 78.3%; Score 36; DB 13; Length 148;
 Best Local Similarity 75.0%; Pred. No. 4.1;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 11:1111
 DB 87 CNDESLIC 94

RESULT 9
 O77643 PRELIMINARY; PRT; 190 AA.
 AC O77643:

DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR.
 GN VEGF.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA-RAMBOULLIET;
 RA Cheung C.Y., Brace R.A.;
 RT "Ovine vascular endothelial growth factor: Nucleotide sequence and
 expression in fetal tissues."
 RL Growth Factors 0:0-0(1998).
 DR EMBL: AF071015; AAC23608.1; -
 DR HSSP: P15692; IYGH.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF00341; PDGF_1.
 DR ProDom: PD001629; PDGF_1.
 DR SMART: SM00141; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 SO SEQUENCE 190 AA; 22342 MW; 0D5E3B3E5C53E739 CRC64;

Query Match 78.3%; Score 36; DB 6; Length 190;
 Best Local Similarity 75.0%; Pred. No. 5.3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
 11:1111
 DB 86 CNEESLIC 93

RESULT 10
 042572 PRELIMINARY; PRT; 194 AA.
 AC 042572;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR 196.
 GN VEGF.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cleaver O., Tonissen K.F., Saha M.S., Krieg P.A.;
 RT "Neovascularization of the Xenopus embryo."
 RL Dev. Dyn. 0:0-0(1997).
 DR EMBL: AF008594; AAB63680.1; -
 DR HSSP: P15692; IYGH.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF00341; PDGF_1.
 DR ProDom: PD001629; PDGF_1.
 DR SMART: SM00141; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 SO SEQUENCE 194 AA; 22672 MW; 85D7BEC7CEFEF17E CRC64;

Query Match 78.3%; Score 36; DB 13; Length 194;
 Best Local Similarity 75.0%; Pred. No. 5.4;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
 11:1111

DB 87 CNEESLIC 94

RESULT 11
 095710 PRELIMINARY; PRT; 421 AA.
 AC 095710;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE Drosophila melanogaster (Fruit fly).
 GN CG3631.
 DR LD29155P.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farina D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (OCF-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY060387; AAL25426.1; -
 SO SEQUENCE 421 AA; 49247 MW; 3C0175ACCE2393CA CRC64;

Query Match 78.3%; Score 36; DB 5; Length 421;
 Best Local Similarity 62.5%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
 111111
 DB 233 CNEESLIC 240

RESULT 12
 09VFF4 PRELIMINARY; PRT; 502 AA.
 AC 09VFF4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE CG3631 PROTEIN.
 GN CG3631.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glotok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003706; AAF55106.1; -
 DR FLYBASE: FBgn0038268; CG3631.
 SQ SEQUENCE 502 AA; 58280 MW; CC2077C417195B5E CRC64;

Query Match 78.3%; Score 36; DB 5; Length 502;
 Best Local Similarity 62.5%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 DB 314 CNEEETVC 321

RESULT 13
 ID 027764 PRELIMINARY; PRT; 1115 AA.
 AC 027764;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE ATPASE.
 OS YE6.
 GN
 OS Plasmodium yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5861;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91161669; PubMed-2150071;
 RA Murakami K., Tanabe K., Takada S.;
 RT "Structure of a Plasmodium yoelii gene-encoded protein homologous to
 RT the Ca(2+)-ATPase of rabbit skeletal muscle sarcoplasmic reticulum.";
 RL J. Cell Sci. 97:487-495(1990).
 DR EMBL: X55197; CAA38982.1; -
 DR HSSP: P04191; 1EUL.
 DR InterPro: IPR004014; Cation_ATPase.
 DR InterPro: IPR001757; E1-E2_ATPase.
 DR InterPro: IPR001454; Hydrolyase.
 DR InterPro: IPR000661; Na_H_K_ATPase.
 DR Pfam: PF00689; Cation_ATPase_C; 1.
 DR Pfam: PF00122; E1-E2_ATPase_N; 1.
 DR Pfam: PF00702; Hydrolyase; 1.
 DR PRINTS: PR00119; CATATPASE.
 DR PROSITE: PS00154; ATPASE_E1_E2; UNKNOWN_1.
 SQ SEQUENCE 1115 AA; 126718 MW; 6B969B7730BE49B6 CRC64;

Query Match 78.3%; Score 36; DB 5; Length 1115;
 Best Local Similarity 62.5%; Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 DB 486 CNEASTIC 493

RESULT 14
 ID 096J14 PRELIMINARY; PRT; 1341 AA.
 AC 096J14;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE KIAA1843 PROTEIN (FRAGMENT).
 GN KIAA1843.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE-BRAIN:
 RC MEDLINE-21245130; PubMed-11347906;
 RX Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. XX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 8:85-95(2001).
 DR EMBL: AB058746; BAB47472.1; -
 DR NON_TER
 FT
 SQ SEQUENCE 1341 AA; 152041 MW; EDE95CEAA624D715 CRC64;

Query Match 78.3%; Score 36; DB 4; Length 1341;
 Best Local Similarity 62.5%; Pred. No. 37;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 DB 206 CNDQSFIC 213

RESULT 15
 ID P78796 PRELIMINARY; PRT; 351 AA.
 AC P78796;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE UNKNOWN PROTEIN (FRAGMENT).
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PR745;
 RX MEDLINE-98162722; PubMed-9501991;
 RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
 RT "Identification of open reading frames in *Schizosaccharomyces pombe*
 RT cDNAs.";
 RL DNA Res. 4:363-369(1997).
 DR EMBL: D89145; BAA13807.1; -
 DR NON_TER
 FT
 SQ SEQUENCE 351 AA; 39291 MW; 02196FDDC8D5A6E CRC64;

Query Match 75.0%; Score 34.5; DB 3; Length 351;
 Best Local Similarity 72.7%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Db 321 CNEESLTDAIC 331

Search completed: June 17, 2002, 16:23:35
Job time: 1609 sec

DR WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues.

PS Claim 49; Page 32; 102pp; English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the exposed loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

XX Sequence 8 AA:

Query Match 100.0%; Score 46; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
| | | | | | | |
Db 1 cneeslic 8

RESULT 2
AAU04528
ID AAU04528 standard; Protein; 9 AA.
XX

AC AAU04528;

XX 26-SEP-2001 (first entry)

VEGF based bicyclic dimeric peptide #2.

XX Human; VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 2..9

FT Disulfide-bond 1 /note- "This bond cyclises the peptide"

FT Disulfide-bond 1 /note- "A disulfide bond forms between residue
1 and residue 1 of an identical peptide to form
a dimeric peptide, or between residue 1 and residue

FT 17 of the sequence appearing as AAU04527 also forming
a dimeric peptide"

XX WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

XX 16-MAY-2000; 2000US-0204590.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Achen MG, Hughes RA, Stackler S, Cendron A;

XX WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues.

PS Claim 59; Page 32; 102pp; English.

XX The sequence represents a dimeric bicyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the exposed loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

XX Sequence 9 AA:

Query Match 100.0%; Score 46; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
| | | | | | | |
Db 2 cneeslic 9

RESULT 3

AAU04522
ID AAU04522 standard; Protein; 73 AA.

XX AAU04522;

XX 26-SEP-2001 (first entry)

DE Human VEGF-D amino acids Val101-Thr 173.
 XX
 XX Human; VEGF-D; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation.
 XX
 XX Homo sapiens.
 XX
 XX WO200152875-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 18-JAN-2001; 2001WO-US01533.
 XX
 XX 18-JAN-2000; 2000US-0176293.
 XX
 XX 16-MAY-2000; 2000US-0204590.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Achen MG, Hughes RA, Stacker S, Cendron A;
 PI
 XX WPI; 2001-442248/47.
 XX
 XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 XX Example 1; Page 90-91; 102pp; English.
 PS
 XX The sequence represents Human VEGF-D (vascular endothelial growth factor)
 CC amino acids Val101-Thr 173, used together with the C-terminal 23
 CC residues of VEGF to make a hybrid theoretical molecule for 3
 CC dimensional modelling. The sequence is used in a method of producing
 CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
 CC distances on opposite antiparallel strands of a peptide loop fragment
 CC from an exposed loop of a growth factor protein and cyclising the peptide
 CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
 CC bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic
 CC peptide with at least one amino acid deleted prior to cyclisation are
 CC used to interfere with angiogenesis, neovascularisation or
 CC lymphangiogenesis in a mammal with a condition characterised by
 CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
 CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
 CC malignant or benign tumour, post-recovery cerebrovascular accident,
 CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
 CC neovascularisation of the liver, excessive hormone-related angiogenic
 CC dysfunction, diabetes induced neovascular sequelae, hypertension induced
 CC neovascular sequelae, or chronic liver infection. The peptides are also
 CC used to modulate vascular permeability in a mammal (the mammal has a
 CC condition characterised by fluid accumulation in peripheral limbs or in
 CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
 CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
 CC peptides are used to interfere with at least one biological activity
 CC induced by VEGF, VEGF-C or -D and are also used in combination with an
 CC anti-inflammatory agent, to treat a chronic inflammation, especially
 CC rheumatoid arthritis, psoriasis and diabetic retinopathy.
 CC
 XX Sequence 73 AA;
 SQ

Query Match 100.0%; Score 46; DB 22; Length 73;
 Best Local Similarity 100.0%; Pred. No. 0.49;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 |||||
 Db 46 cneeslic 53

RESULT 4
 AAU04520

ID AAU04520 standard; protein; 96 AA.
 XX
 XX AAU04520:
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX Human VEGF-D amino acids Val101-PRO186.
 DE
 XX
 XX Human; VEGF-D; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation.
 XX
 XX Homo sapiens.
 XX
 XX WO200152875-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 18-JAN-2001; 2001WO-US01533.
 XX
 XX 18-JAN-2000; 2000US-0176293.
 XX
 XX 16-MAY-2000; 2000US-0204590.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Achen MG, Hughes RA, Stacker S, Cendron A;
 PI
 XX WPI; 2001-442248/47.
 XX
 XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 XX Example 1; Page 89; 102pp; English.
 PS
 XX The sequence represents Human VEGF-D (vascular endothelial growth factor)
 CC amino acids Val101-PRO186. The sequence is used in a method of producing
 CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
 CC distances on opposite antiparallel strands of a peptide loop fragment
 CC from an exposed loop of a growth factor protein and cyclising the peptide
 CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
 CC bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic
 CC peptide with at least one amino acid deleted prior to cyclisation are
 CC used to interfere with angiogenesis, neovascularisation or
 CC lymphangiogenesis in a mammal with a condition characterised by
 CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
 CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
 CC malignant or benign tumour, post-recovery cerebrovascular accident,
 CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
 CC neovascularisation of the liver, excessive hormone-related angiogenic
 CC dysfunction, diabetes induced neovascular sequelae, hypertension induced
 CC neovascular sequelae, or chronic liver infection. The peptides are also
 CC used to modulate vascular permeability in a mammal (the mammal has a
 CC condition characterised by fluid accumulation in peripheral limbs or in
 CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
 CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
 CC peptides are used to interfere with at least one biological activity
 CC induced by VEGF, VEGF-C or -D and are also used in combination with an
 CC anti-inflammatory agent, to treat a chronic inflammation, especially
 CC rheumatoid arthritis, psoriasis and diabetic retinopathy.
 CC
 XX Sequence 96 AA;
 SQ

Query Match 100.0%; Score 46; DB 22; Length 96;
 Best Local Similarity 100.0%; Pred. No. 0.64;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
 |||||
 Db 46 cneeslic 53

RESULT 5
AAV23889
ID AAV23889 standard; Protein; 109 AA.
XX
AC AAV23889;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human vascular endothelial growth factor (VEGF)-D.
XX
KW Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
KW tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
KW wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
XX
OS Homo sapiens.
XX
PN WO9933485-A1.
XX
PD 08-JUL-1999.
XX
PE 23-DEC-1998; 98WO-US27373.
XX
PR 29-MAY-1998; 98US-0087392.
PR 24-DEC-1997; 97AU-0001131.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Altalo K, Stacker SA;
XX
DR WPI; 1999-405368/34.
XX
PT A human cell line stably expressing vascular endothelial growth
XX factor D, useful for treating melanomas or tumours expressing VEGF-D
XX
PS Claim 6; Page 72; 79pp; English.
XX
CC The present sequence represents human vascular endothelial growth
CC factor (VEGF)-D. The specification describes a human cell line
CC which stably expresses VEGF-D, or fragments/analogs having VEGF-D
CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
CC antibodies), are useful for the treatment or alleviation of malignant
CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
XX
SQ Sequence 109 AA;
XX
Query Match 100.0%; Score 46; DB 20; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNEESLIC 8
DB 54 cneeslic 61
XX
RESULT 6
AAB11931
ID AAB11931 standard; Protein; 109 AA.
XX

AC AAB11931;
XX
DT 20-NOV-2000 (first entry)
XX
DE Human truncated VEGF-D.
XX
KW Truncated VEGF-D; vascular endothelial growth factor; human;
KW monoclonal antibody; VEGF receptor; VEGFR-2; VEGFR-3;
KW vascular permeability disorder; endothelial cell proliferative disorder;
KW angiogenic disorder; lymphangiogenic disorder;
KW neovascularisation disorder; endothelial cell differentiation disorder;
KW cancer; diabetic retinopathy; psoriasis; arthropathy; pulmonary oedema;
KW detection; diagnosis; imaging; lymphatic vasculature.
XX
OS Homo sapiens.
XX
PN WO200037025-A2.
XX
PD 29-JUN-2000.
XX
PE 21-DEC-1999; 99WO-US31332.
XX
PR 21-DEC-1998; 98US-0113254.
PR 17-MAY-1999; 99US-0134556.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Stacker SA;
XX
DR WPI; 2000-442498/38.
XX
PT Novel compositions comprising antibodies reactive to vascular
XX endothelial growth factor-D, useful for treating, e.g. angiogenesis,
XX lymphangiogenesis and neovascularization disorders.
XX
PS Claim 1; Fig 1; 44pp; English.
XX
CC This sequence represents a 109 amino acid truncated human VEGF-D
CC (vascular endothelial growth factor D), lacking both the N- and
CC C-terminal regions. The invention relates to a monoclonal antibody, or
CC fragments thereof, which is specifically reactive with the truncated
CC human VEGF-D, and methods of preparing the antibody. The antibody of the
CC invention interferes with the binding of VEGF-D to the VEGF receptors
CC VEGFR-2 and VEGFR-3, but does not interfere with the binding of VEGF to
CC these receptors and additionally is not reactive with VEGF-C. The
CC antibody may be used to treat disorders associated with vascular
CC permeability, endothelial cell proliferation, angiogenesis,
CC lymphangiogenesis, neovascularisation and endothelial cell
CC differentiation, especially cancer, diabetic retinopathy, psoriasis, and
CC arthropathies. The antibody may also be used to treat fluid accumulation
CC in the heart and/or lung via modulation of vascular permeability. It may
CC additionally be used to detect VEGF-D and may be used to image lymphatic
CC vasculature in tissue.
XX
SQ Sequence 109 AA;
XX
Query Match 100.0%; Score 46; DB 21; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNEESLIC 8
DB 54 cneeslic 61
XX
RESULT 7
AAW53240
ID AAW53240 standard; Protein; 325 AA.
XX
AC AAW53240;
XX
DT 03-AUG-1998 (first entry)

XX Sequence 325 AA;
SQ
Query Match 100.0%; Score 46; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNEESLIC 8
|||||||
Db 117 cneeslic 124
RESULT 9
AAW49036
ID AAW49036 standard; Protein: 354 AA.
XX
AC AAW49036;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human zvegf2 growth factor.
XX
KW Human zvegf2 growth factor; mitogen; fibroblast; smooth muscle cell;
KW venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
KW angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
KW scleroderma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Peptide /note="Signal peptide"
FT Peptide 24..108
FT Binding-site /note="Pro-region"
FT Binding-site 109..197
FT Region /note="Receptor binding domain"
FT Region 206..256
FT Region /note="Cysteine-rich domain"
FT Region 257..274
FT Region /note="Balblanl ring motif"
FT Region 275..294
FT Region /note="Balblanl ring motif"
FT Region 295..354
FT Region /note="Cysteine-rich domain"
XX
PN WO9824811-A2.
XX
PD 11-JUN-1998.
XX
PE 20-NOV-1997; 97WO-US20888.
XX
PR 18-SEP-1997; 97US-0933455.
PR 06-DEC-1996; 96US-0759657.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
XX
DR WPI: 1998-333256/29.
DR N-PSDB: AAV32823.
XX
PT New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX
XX Claim 1; Pages 53-54; 77pp; English.
XX
CC The present sequence represents a human zvegf2 growth factor encoded
CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
CC revascularisation of tissue or the re-endothelialisation of vascular

CC tissue. zvegf2 is particularly claimed to be useful for the treatment
CC of full-thickness skin wounds, including venous stasis ulcers and
CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
CC additive in tissue adhesives for promoting revascularisation of the
CC healing tissue. Antagonists against zvegf2 can be used to block its
CC mitogenic, chemotactic and angiogenic effects. The antagonists may
CC therefore be useful for reducing growth of solid tumours by inhibiting
CC neovascularisation of the developing growth of solid tumours by inhibiting
CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
CC arthritis, and scleroderma.
XX
SQ Sequence 354 AA;
QY 1 CNEESLIC 8
|||||||
Db 146 cneeslic 153
RESULT 10
AAW53241
ID AAW53241 standard; Protein: 354 AA.
XX
AC AAW53241;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Homo sapiens.
XX
FH Homo sapiens.
FT WO9807832-A1.
XX
PN 26-FEB-1998.
XX
PD 21-AUG-1997; 97WO-US14696.
XX
PE 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Achen MG, Aitalo K, Stackner SA, Wilks AF;
XX
DR WPI: 1998-179057/16.
DR N-PSDB: AAV20807.
XX
PT New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
XX Claim 16; Pages 60-61; 101pp; English.

Query Match 100.0%; Score 46; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
DB 146 cneeslic 153

RESULT 13

AAB29049
ID AAB29049 standard; Protein: 354 AA.

XX AAB29049;

DT 31-JAN-2001 (first entry)

DE Human VEGF-D protein sequence.

XX Human; Flt4; fms-like tyrosine kinase 4; lymphoedema;
XX vascular endothelial growth factor receptor 3; VEGFR-3;
KM Mlroy-Nonne syndrome; lymphoedema praecox; VEGF-D;
XX vascular endothelial growth factor D.

OS Homo sapiens.

PN WO200058511-A1.

PD 05-OCT-2000.

PF 26-MAR-1999; 99WO-US06133.

PR 26-MAR-1999; 99WO-US06133.

PA (LUDW-) LUDWIG INST CANCER RES.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.

PI Ferrell RE, Alltalo K, Finegold DN, Karkkainen M;

DR WPI: 2000-679298/66.

XX N-PSDB: AAC62407.

PT Screening a human subject for increased risk of developing a lymphatic
PT disorder, comprises assaying a nucleic acid to determine a mutation
PT altering the sequence of a vascular endothelial growth factor
PT receptor-3.

PS Disclosure: Page 64-65; 76pp; English.

CC The present sequence is the protein sequence of the human vascular
CC endothelial growth factor D (VEGF-D). It was used to demonstrate the
CC methods of the invention, which involve the screening of individuals to
CC determine which vascular endothelial growth factor receptor 3 (VEGFR-3,
CC also known as Flt4 or fms-like tyrosine kinase 4) alleles they possess
CC and thus their likelihood of developing hereditary lymphoedema.
CC Conditions associated with lymphoedema include Mlroy-Nonne syndrome,
CC which is early onset lymphoedema and lymphoedema praecox, which is late
CC onset.

XX Sequence 354 AA;

Query Match 100.0%; Score 46; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
DB 146 cneeslic 153

RESULT 14

AAV70750
ID AAV70750 standard; Protein: 354 AA.

XX AAV70750;

DT 17-AUG-2000 (first entry)

DE Human prepro-vascular endothelial growth factor D.

XX Human; receptor tyrosine kinase; RTK; Flt4; fms-like tyrosine kinase 4;
XX VEGFR-3; vascular endothelial growth factor receptor 3; chromosome 5q35;
XX cytosolic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
KM neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
XX sarcoma; malignancy; VEGF-D; vascular endothelial growth factor D.

OS Homo sapiens.

PN WO200021560-A1.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23525.

PR 09-OCT-1998; 98US-0169079.

PA (LUDW-) LUDWIG INST CANCER RES.

PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.

PI Alltalo K, Kaipainen A, Valtola R, Jussila L;

DR WPI: 2000-317850/27.

PT Treating neoplastic diseases such as lymphoma, carcinomas, melanomas
PT and sarcomas, involves administering a compound capable of inhibiting
PT binding of ligand proteins to fms-like tyrosine kinase-4 receptor
XX Example 15-17; Page 142-143; 148pp; English.

CC The patent discloses a method to treat neoplastic disease characterised
CC by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also
CC referred as vascular endothelial growth factor receptor-3, VEGFR-3) in
CC endothelial cells of blood vessels adjacent to malignant neoplasm. The
CC method involves administering a compound that inhibits binding of a
CC ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
CC endothelial cells. The compound is useful for treating neoplastic disease
CC such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
CC and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used
CC for manufacturing medicament useful for diagnostic screening, imaging and
CC treatment of malignancies characterised by Flt4-expressing blood cells.
CC The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
CC and 4.5 kb mRNAs which differ in their 3' sequences and are
CC differentially expressed in HEL and DAPI cell lines. Flt4
CC belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
CC It is used as a target for tumour imaging and anti-tumour therapy.
CC The present sequence is a human prepro-vascular endothelial growth
CC factor D (VEGF-D), a specific example of Flt4 binding compound. A
CC recombinantly matured VEGF-D lacking residues 1-92 and 202-354
CC retains the ability to activate VEGFR-2 and VEGFR-3 receptors and
CC associate as non-covalently linked dimers.

XX Sequence 354 AA;

Query Match 100.0%; Score 46; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
DB 146 cneeslic 153

RESULT 15
AAV70983

ID AAV70983 standard; Protein: 354 AA.

AC AAV70983;

DT 09-AUG-2000 (first entry)

DE Human vascular endothelial growth factor (VEGF)-D protein.

XX Vascular endothelial growth factor-D: VEGF; human; re-endothelialisation;

KW vascular endothelial growth factor receptor; VEGFR; vascular trauma;

KW blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;

KW restenosis; stenosis; percutaneous transluminal coronary angioplasty.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OY 1 CNEESTLIC 8
Db 146 cneestlic 153

Search completed: June 17, 2002, 16:02:12
Job time: 421 sec

Location/Qualifiers
1..21
/label=Signal_peptide
/note="Cleavage results in partially-processed VEGF-D protein"
22..92
/label=Amino_terminal_peptide
/note="Cleavage results in fully-processed mature VEGF-D protein"
93..201
/label=Recombinantly_matured_VEGF_D_protein
/note="Processed vascular epithelial growth factor-D"
202..354
/label=Carboxy_terminal_peptide
/note="Cleavage results in partially-processed VEGF-D protein"

MO200024412-A2.

04-MAY-2000.

26-OCT-1999; 99WO-US24054.

26-OCT-1998; 98US-0105587.

(LUDWIG INST CANCER RES.
(UYHE-) UNIV HELSINKI LICENSING LTD OY.
(YIAH/) YIA-HERTTUALA S.

Yia-herTTuala S, Alltalo K, Hiltunen MO, Jeltsch MM, Achen MG;

WPI: 2000-350584/30.

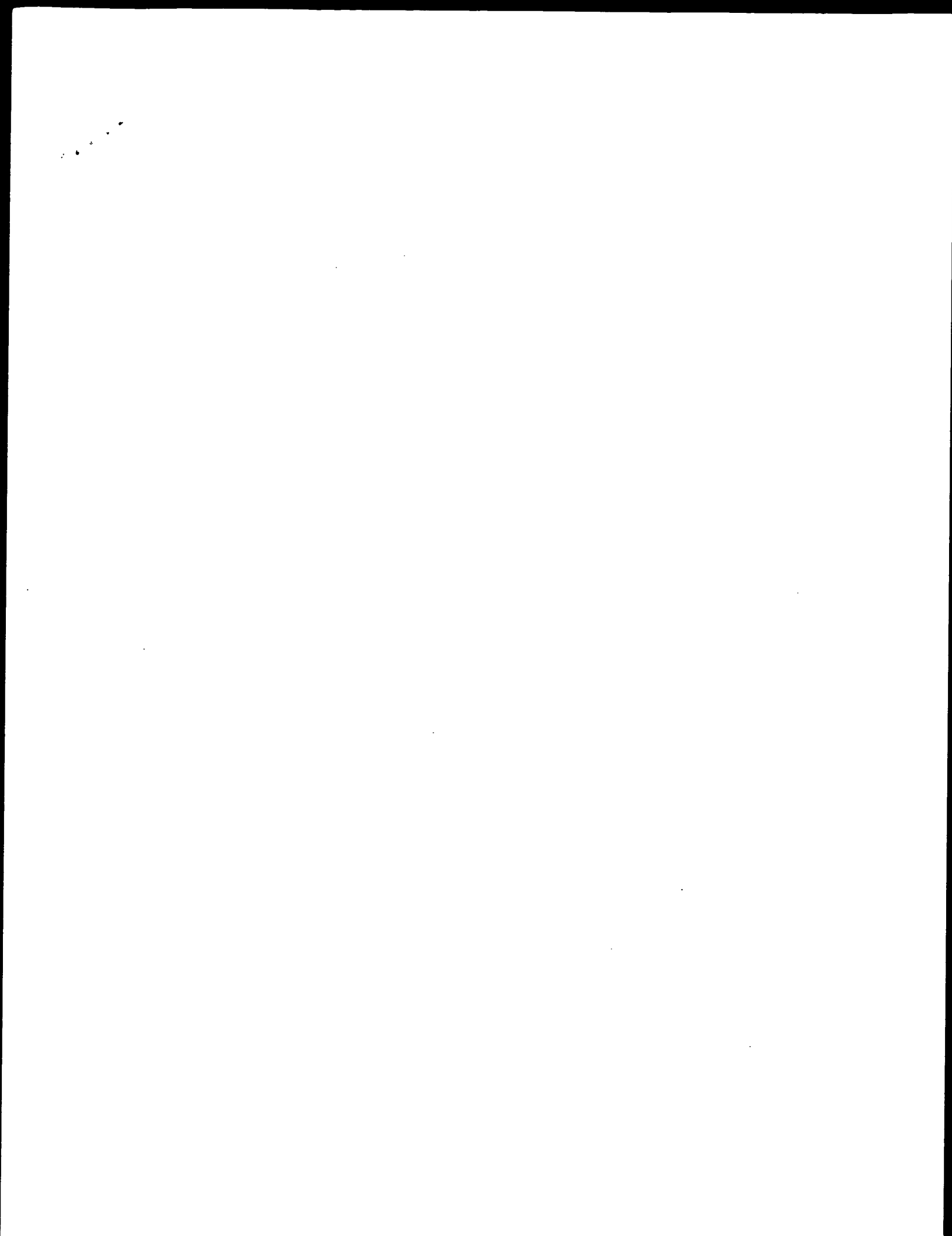
N-PSDB; AAD00340.

Preventing stenosis and restenosis in mammals using vascular endothelial growth factor proteins or the nucleic acids encoding them -
Disclosure: Page 53-55; 61pp; English.

The present amino acid sequence is the complete human prepro-vascular endothelial growth factor (VEGF)-D. VEGF-D has the ability to stimulate re-endothelialisation of an injured blood vessel, without significant stimulation of smooth muscle cell proliferation. It can bind to and stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or VEGFR-3 phosphorylation in cells that express such receptors. An anti-restenosis agent comprising either a VEGF-D gene or protein is used in a method to reduce or prevent restenosis and stenosis of a blood vessel following vascular trauma e.g., cardiovascular surgery and percutaneous transluminal coronary angioplasty.

Sequence 354 AA:

Query Match 100.0%; Score 46; DB 21; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.3; Mismatches 0; Gaps 0;
Matches 8; Conservative 0; Indels 0;



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:35 ; Search time 75.98 Seconds
(without alignments)
2.572 Million cell updates/sec

Title: US-09-761-636A-6
Perfect score: 46
Sequence: 1 CNESSLIC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfilest1.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	46	100.0	325 4 US-08-915-795-3	Sequence 3, Appli
2	46	100.0	354 4 US-08-915-795-5	Sequence 3, Appli
3	36	78.3	120 6 5194596-9	Patent No. 5194596
4	36	78.3	120 6 5219739-9	Patent No. 5219739
5	36	78.3	164 6 5194596-17	Patent No. 5194596
6	36	78.3	164 6 5219739-17	Patent No. 5219739
7	36	78.3	164 6 5219739-18	Patent No. 5219739
8	36	78.3	190 6 5332671-3	Patent No. 5332671
9	36	78.3	321 4 US-08-915-795-9	Sequence 9, Appli
10	36	78.3	358 4 US-08-915-795-8	Sequence 9, Appli
11	33	71.7	146 4 US-08-586-039B-33	Sequence 33, Appli
12	33	71.7	189 1 US-08-469-427A-15	Sequence 15, Appli
13	33	71.7	190 2 US-08-589-063C-20	Sequence 20, Appli
14	33	71.7	190 4 US-08-586-039B-31	Sequence 31, Appli
15	33	71.7	214 4 US-08-586-039B-35	Sequence 35, Appli
16	33	71.7	278 2 US-08-460-309-13	Sequence 13, Appli
17	33	71.7	278 2 US-08-125-077-13	Sequence 13, Appli
18	33	71.7	279 1 US-08-152-019A-29	Sequence 29, Appli
19	33	71.7	1196 1 US-08-144-121-4	Sequence 4, Appli
20	33	71.7	1196 2 US-08-735-893-4	Sequence 43, Appli
21	32	69.6	12 3 US-08-742-243-43	Sequence 43, Appli
22	32	69.6	12 3 US-08-742-243-44	Sequence 44, Appli
23	32	69.6	12 3 US-08-742-243-45	Sequence 45, Appli
24	32	69.6	109 3 US-08-691-794-3	Sequence 3, Appli
25	32	69.6	110 4 US-09-392-932-11	Sequence 11, Appli
26	32	69.6	121 6 5194596-19	Patent No. 5194596
27	32	69.6	121 6 5219739-20	Patent No. 5219739

28	32	69.6	145 3 US-08-784-551C-2	Sequence 2, Appli
29	32	69.6	145 4 US-09-392-932-2	Sequence 2, Appli
30	32	69.6	147 3 US-08-807-992B-1	Sequence 1, Appli
31	32	69.6	147 4 US-09-392-932-1	Sequence 1, Appli
32	32	69.6	165 6 5194596-18	Patent No. 5194596
33	32	69.6	165 6 5219739-19	Patent No. 5219739
34	32	69.6	191 3 US-08-567-200A-2	Sequence 2, Appli
35	32	69.6	191 3 US-08-807-992B-2	Sequence 2, Appli
36	32	69.6	191 3 US-08-691-794-2	Sequence 3, Appli
37	32	69.6	191 4 US-08-795-430-56	Sequence 56, Appli
38	32	69.6	191 4 US-09-392-932-3	Sequence 3, Appli
39	32	69.6	191 6 5332671-4	Patent No. 5332671
40	32	69.6	214 6 5240848-11	Patent No. 5240848
41	32	69.6	215 3 US-08-807-992B-3	Sequence 3, Appli
42	32	69.6	215 4 US-08-586-039B-49	Sequence 49, Appli
43	32	69.6	215 6 5219739-22	Patent No. 5219739
44	32	69.6	215 6 5240848-7	Patent No. 5240848
45	32	69.6	231 5 PCT-US96-09001-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

APPLICANT: Karl ALITALO

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKewen, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE: 536

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 325 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

TISSUE TYPE: Human Breast

US-08-915-795-3

Query Match 100.0%; Score 46; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNESSLIC 8
Db 117 CNESSLIC 124

RESULT 2

US-08-915-795-5

; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESS: Evenson, McKown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 354 amino acids

TYPE: amino acid

STRANDEDNESS: single

MOLECULE TYPE: linear

HYPOTHETICAL: NO

ORIGINAL SOURCE:

TISSUE TYPE: Human Lung

US-08-915-795-5

Query Match 100.0%; Score 46; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CNESSLIC 8

Db 146 CNESSLIC 153

RESULT 3

5194596-9

Patent No. 5194596

APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN

C.; MITCHELL, RICHARD L.

TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL

GROWTH FACTOR

NUMBER OF SEQUENCES: 32

CURRENT APPLICATION DATA:

4

APPLICATION NUMBER: US/07/450,883

FILING DATE: 14-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 387,545

FILING DATE: 27-JUL-1989

SEQ ID NO: 9

LENGTH: 120

5194596-9

Query Match 78.3%; Score 36; DB 6; Length 120;
Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNESSLIC 8

Db 60 CNESSLIC 67

RESULT 4

5219739-9

Patent No. 5219739

APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN C.; MITCHELL, RICHARD L.

TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND

HVEGF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN

VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVEGF121

NUMBER OF SEQUENCES: 40

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/559,041

FILING DATE: 27-JUL-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 450,883

FILING DATE: 14-DEC-1989

APPLICATION NUMBER: 387,545

FILING DATE: 27-JUL-1989

SEQ ID NO: 9

LENGTH: 120

5219739-9

Query Match 78.3%; Score 36; DB 6; Length 120;
Best Local Similarity 75.0%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNESSLIC 8

Db 60 CNESSLIC 67

RESULT 5

5194596-17

Patent No. 5194596

APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN

C.; MITCHELL, RICHARD L.

TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL

GROWTH FACTOR

NUMBER OF SEQUENCES: 32

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/450,883

FILING DATE: 14-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 387,545

FILING DATE: 27-JUL-1989

SEQ ID NO: 17

LENGTH: 164

5194596-17

Query Match 78.3%; Score 36; DB 6; Length 164;
Best Local Similarity 75.0%; Pred. No. 16;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
11:1111
Db 60 CNDSESLC 67

RESULT 6
5219739-17
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; HVEGF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVEGF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:17:
; LENGTH: 164
5219739-17

Query Match 78.3%; Score 36; DB 6; Length 164;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
11:1111
Db 60 CNDSESLC 67

RESULT 7
5219739-18
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; HVEGF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVEGF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:18:
; LENGTH: 164
5219739-18

Query Match 78.3%; Score 36; DB 6; Length 164;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
11:1111
Db 60 CNDSESLC 67

RESULT 8
5332671-3
; Patent No. 5332671
; APPLICANT: FERRARA, NAPOLEONE; LEUNG, DAVID W.H.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR AND DNA ENCODING SAME

NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/389,722
FILING DATE: 04-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 369,424
FILING DATE: 21-JUN-1989
APPLICATION NUMBER: 351,117
FILING DATE: 12-MAY-1989
SEQ ID NO:3:
LENGTH: 190
5332671-3

Query Match 78.3%; Score 36; DB 6; Length 190;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
11:1111
Db 86 CNDSESLC 93

RESULT 9
US-08-915-795-9
; Sequence 9, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: MARC G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; TISSUE TYPE: Mouse Lung
; US-08-915-795-9

Query Match 78.3%; Score 36; DB 4; Length 321;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
Db 146 CNEGVWC 153

RESULT 10
US-08-915-795-8

; Sequence 8, Application US/08915795
; Patent No. 6235713

; GENERAL INFORMATION:

; APPLICANT: Marc G. ACHEN

; APPLICANT: Andrew F. WILKS

; APPLICANT: Steven A. STACKER

; APPLICANT: Kari ALITALO

; TITLE OF INVENTION: GROWTH FACTOR

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

; STREET: 1200 G Street, NW, Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: United States of America

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,795

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 1064/42983

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; TELEX: N/A

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 358 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; TISSUE TYPE: Mouse Lung

; US-08-915-795-8

Query Match

Best Local Similarity 78.3%; Score 36; DB 4; Length 358;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 151 CNEGVWC 158

OY 1 CNEESLIC 8

Db 151 CNEGVWC 158

RESULT 11

US-08-586-039B-33

; Sequence 33, Application US/08586039B

; Patent No. 6140073

; GENERAL INFORMATION:

; APPLICANT: Bayne, Marvin L.

; APPLICANT: Thomas Jr., Kenneth A.

; TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 E. Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Microsoft Word 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/586,039B

FILING DATE: 16-JAN-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/124,259

FILING DATE: 20-SEP-1993

APPLICATION NUMBER: 07/676,436

FILING DATE: 28-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hand, J. Mark

REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 18361DA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3905

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 146 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-586-039B-33

Query Match

Best Local Similarity 71.7%; Score 33; DB 4; Length 146;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 86 CNEDELEC 93

OY 1 CNEESLIC 8

Db 86 CNEDELEC 93

RESULT 12

US-08-469-427A-15

; Sequence 15, Application US/08469427A

; Patent No. 5607918

; GENERAL INFORMATION:

; APPLICANT: Eriksson, Ulf

; APPLICANT: Olofsson, Birgitte

; APPLICANT: Alitalo, Kari

; APPLICANT: Pajusola, Katri

; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,427A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,651
FILING DATE: 01-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 41979cp2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-427A-15

Query Match 71.7%; Score 33; DB 1; Length 189;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
DB 85 CNDEALEC 92

RESULT 13

US-08-569-063C-20
Sequence 20, Application US/08569063C
Patent No. 5928939

GENERAL INFORMATION:

APPLICANT: ERIKSSON, Ulf
APPLICANT: OLOFSSON, Birgitta

APPLICANT: ALITALO, Kari

APPLICANT: PATUSOLA, Katari

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
TITLE OF INVENTION: DNA CODING THEREFOR

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan, P.L.L.C.

STREET: 1200 G Street, N.W., Suite 700
City: Washington

State: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/569,063C

FILING DATE: 06-DEC-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/469,427

FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/397,651

FILING DATE: 01-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Evans, Joseph D

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/41979CP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 190 amino acids

TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-063C-20

Query Match 71.7%; Score 33; DB 2; Length 190;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
DB 86 CNDEALEC 93

RESULT 14

US-08-586-039B-31
Sequence 31, Application US/08586039B
Patent No. 6140073

GENERAL INFORMATION:

APPLICANT: Bayne, Marvin L.

APPLICANT: Thomas Jr., Kenneth A.

TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C

TITLE OF INVENTION: SUBUNIT

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 E. Lincoln Avenue

City: Rahway

State: New Jersey

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Microsoft Word 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/586,039B

FILING DATE: 16-JAN-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/124,259

FILING DATE: 20-SEP-1993

APPLICATION NUMBER: 07/676,436

FILING DATE: 28-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hand, J. Mark

REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 18361DA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-4720

TELEFAX: (908) 594-3905

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 190 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-586-039B-31

Query Match 71.7%; Score 33; DB 4; Length 190;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 8
DB 86 CNDEALEC 93

RESULT 15

US-08-586-039B-35
; Sequence 35, Application US/08586039B
; Patent No. 6140073
; GENERAL INFORMATION:
; APPLICANT: Bayne, Marvin L.
; APPLICANT: Thomas Jr., Kenneth A.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C
; TITLE OF INVENTION: SUBUNIT
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 E. Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,039B
; FILING DATE: 16-JAN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/124,259
; FILING DATE: 20-SEP-1993
; APPLICATION NUMBER: 07/676,436
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 18361DA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-586-039B-35

Query Match 71.7% Score 33; DB 4; Length 214;
Best Local Similarity 62.5%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 8
||:|:|
Db 86 CNDEALEC 93

Search completed: June 17, 2002, 15:56:36
Job time: 85 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 17, 2002, 15:58:20 ; Search time 95.43 Seconds
(without alignments)
11.076 Million cell updates/sec

Title: US-09-761-636a-7
Perfect score: 61
Sequence: 1 CISVPLTSVPC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	39	63.9	151	2 P00506	hypothetical prote
2	39	63.9	169	2 A72466	hypothetical prote
3	39	63.9	207	2 I47061	collagenase inhibi
4	39	63.9	925	2 T29311	hypothetical prote
5	39	63.9	1062	2 T46444	hypothetical prote
6	38	62.3	324	2 T31992	hypothetical prote
7	38	62.3	503	2 S62018	probable membrane
8	38	62.3	616	2 T29234	hypothetical prote
9	38	62.3	1829	2 T34239	hypothetical prote
10	37	60.7	184	2 B96814	hypothetical prote
11	37	60.7	527	2 H85135	hypothetical prote
12	37	60.7	583	2 F70592	probable 1pqb pro
13	37	60.7	788	2 G64707	cation-transportin
14	37	60.7	788	2 E71813	probable component
15	37	60.7	1787	2 B72727	hypothetical prote
16	36	59.0	165	2 JY0046	hypothetical 18k p
17	36	59.0	174	2 T16392	hypothetical prote
18	36	59.0	261	2 T46975	lysine-tRNA ligas
19	36	59.0	533	2 C64367	hypothetical prote
20	36	59.0	954	2 G71496	hypothetical prote
21	35	57.4	112	2 E72785	phospholipase A2 (
22	35	57.4	119	2 A59047	metalloprotease
23	35	57.4	207	2 A35685	metalloprotease
24	35	57.4	207	2 I46964	hypothetical prote
25	35	57.4	296	2 T34460	4-hydroxybenzoate
26	35	57.4	302	2 B81696	hypothetical prote
27	35	57.4	308	2 T29756	hypothetical prote
28	35	57.4	342	2 T25143	hypothetical prote
29	35	57.4	342	2 T25143	hypothetical prote

30	35	57.4	354	2 T31861	hypothetical prote
31	35	57.4	363	2 T02522	origin recognition
32	35	57.4	403	2 C75405	streptomycin biosy
33	35	57.4	410	2 T18995	hypothetical prote
34	35	57.4	457	2 B84725	probable glucosylt
35	35	57.4	537	2 T48599	hypothetical prote
36	35	57.4	756	2 T20109	hypothetical prote
37	35	57.4	790	2 T25095	hypothetical prote
38	35	57.4	1099	2 T16283	genome polyprotein
39	35	57.4	1014	1 JCS620	hypothetical prote
40	35	56.6	1246	2 T00826	hypothetical prote
41	34.5	56.6	1816	2 A84845	probable ABC trans
42	34	55.7	61	2 H90901	probable lipoprote
43	34	55.7	76	2 C90971	probable lipoprote
44	34	55.7	76	2 JE0003	hypothetical 8.6k
45	34	55.7	112	2 AE0663	probable multitrng

ALIGNMENTS

RESULT 1
P00506
hypothetical protein 151 - fowlpox virus (fragment)
N:Alternate names: ORF3 protein
C:Species: fowlpox virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 03-Nov-2000
C:Accession: P00506; S27935
R:Ogawa, R.; Calvert, J.G.; Yanagida, N.; Nazerian, K.
J. Gen. Virol. 74, 55-64, 1993
A:Title: Insertional inactivation of a fowlpox virus homologue of the vaccinia v
A:Reference number: J01894; MUID:93138784
A:Accession: P00506
A:Molecule type: DNA
A:Residues: 1-151 <OGA>
A:Cross-references: GB:M88588; NID:9333522; PIDN:AAA47188.1; PID:9333525
A>Note: Submitted to the EMBL Data Library, May 1992

Query Match 63.9%; Score 39; DB 2; Length 151;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

RESULT 2
A72466
hypothetical protein APE2372 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: A72466
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki,
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon,
A:Reference number: A72450; MUID:99310339
A:Accession: A72466
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <KAN>
A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BA81385.1; PID:01045171;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2372

Query Match 63.9%; Score 39; DB 2; Length 169;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 C1SVPLTSVPC 10
 11 : 11 : 11 :
 Db 57 C1GIPVSVPC 66

RESULT 3

collagenase inhibitor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 16-Jul-1999
 C:Accession: I47061
 R:Tanaka, T.; Andoh, N.; Takeya, T.; Sato, E.
 Moll. Cell. Endocrinol. 83, 65-71, 1992
 A:Title: Differential screening of ovarian cDNA libraries detected the expression of the
 A:Reference number: I47061; MUID:92201478
 A:Accession: I47061
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-207 <TRAN>
 A:Cross-references: GB:S96211; NID:g247729; PIDN:AAB21865.1; PID:g247730
 C:Superfamily: metalloproteinase inhibitor

Query Match 63.9%; Score 39; DB 2; Length 207;
 Best Local Similarity 54.5%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 C1SVPLTSVPC 11
 11 : 11 : 11 :
 Db 150 C1GVPLTSVPC 160

RESULT 4

hypothetical protein F36D4.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29311
 R:Pauley, A.; Gattung, S.
 submitted to the EMBL data library March 1996
 A:Description: The sequence of C. elegans cosmid F36D4.
 A:Reference number: Z20603
 A:Accession: T29311
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-925 <PAU>
 A:Cross-references: EMBL:U53181; PIDN:AAA93485.1; GSPDB:GN00023; CESP:F36D4.3
 A:Experimental source: strain Bristol N2; clone F36D4
 C:Genetics:
 A:Gene: CESP:F36D4.3
 A:Map position: 5
 A:Introns: 24/3; 56/2; 103/3; 187/1; 387/2; 429/3; 455/3; 516/1; 555/1; 782/2; 882/3

Query Match 63.9%; Score 39; DB 2; Length 925;
 Best Local Similarity 50.0%; Pred. No. 64;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ISVPLTSVPC 11
 11 : 11 : 11 :
 Db 2 VSLPMSNVPC 11

RESULT 5

hypothetical protein DKFZp434N1427.1 - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46444
 R:Blocker, H.; Boeher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23032
 A:Accession: T46444

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1062 <AA>
 A:Cross-references: EMBL:AL137701
 A:Experimental source: adult testis; clone DKFZp434N1427
 C:Genetics:
 A:Note: DKFZp434N1427.1

Query Match 63.9%; Score 39; DB 2; Length 1062;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 SVPLTSVPC 11
 11 : 11 : 11 :
 Db 25 S1PLSSTPC 33

RESULT 6

hypothetical protein C49D10.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T31992
 R:Henkhaus, J.; Wohldmann, P.; Beck, C.
 submitted to the EMBL data library, July 1997
 A:Description: The sequence of C. elegans cosmid C49D10.
 A:Reference number: Z21108
 A:Accession: T31992
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-324 <HEND>
 A:Cross-references: EMBL:AF016665; PIDN:AAC71178.1; GSPDB:GN00020; CESP:C49D10.3
 A:Experimental source: strain Bristol N2; clone C49D10
 C:Genetics:
 A:Gene: CESP:C49D10.3
 A:Map position: 2
 A:Introns: 105/1
 C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 62.3%; Score 38; DB 2; Length 324;
 Best Local Similarity 54.5%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 C1SVPLTSVPC 11
 11 : 11 : 11 :
 Db 70 C1GVPLTLPC 80

RESULT 7

probable membrane protein YDR539W - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein D3703.2
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 17-Mar-2000
 C:Accession: S62018
 R:Dieckrich, F.S.; Mulligan, J.; Allen, E.; Aranjó, R.; Aviles, E.; Bero, A.; Carpe
 H.; Lin, D.; Mosedale, D.; Nakahara, K.; Nemath, A.; Oefner, P.; On, C.; Petel, F.X
 W.
 submitted to the EMBL data library, December 1995
 A:Reference number: S62017
 A:Accession: S62018
 A:Molecule type: DNA
 A:Residues: 1-503 <DI>
 A:Cross-references: EMBL:U43834; NID:g1165292; PID:g1165294; GSPDB:GN00004; MIPS:YDI
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: MIPS:YDR539W
 A:Map position: 4R
 C:Superfamily: conserved hypothetical protein s110936
 C:Keywords: transmembrane protein
 F:210-226/Domain: transmembrane #status predicted <TM>

Query Match 62.3%; Score 38; DB 2; Length 503;
 Best Local Similarity 60.0%; Pred. No. 53;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 ISVPLTSPVC 11
 |||||
 DB 115 ITVPSSAPC 124

RESULT 8
 T929234
 hypothetical protein F55G1.13 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T929234
 R:Murray, J.; Le, T.T.
 Submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid F55G1.
 A:Reference number: 220591
 A:Accession: T929234
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-616 <MOR>
 A:Cross-references: EMBL:U58750; PIDN:AB00653.1; GSPDB:GN00022; CESP:F55G1.13
 A:Experimental source: strain Bristol N2; clone F55G1
 C:Genetics:
 A:Gene: CESP:F55G1.13
 A:Map position: 4
 A:introns: 98/1; 136/1; 230/1; 256/1; 363/1; 401/1; 443/1; 484/1; 528/2; 551/3

Query Match 62.3%; Score 38; DB 2; Length 616;
 Best Local Similarity 54.5%; Pred. No. 64;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CISVPLTSPVC 11
 |||||
 DB 362 CVCNPTSPC 372

RESULT 9
 T34239
 hypothetical protein F26F12.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34239
 R:Wilson, R.; Bentley, D.; Gattung, S.
 Submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid F26F12.
 A:Reference number: 221493
 A:Accession: T34239
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1829 <WID>
 A:Cross-references: EMBL:U55373; PIDN:AAC25894.1; GSPDB:GN00023; CESP:F26F12.7
 A:Experimental source: strain Bristol N2; clone F26F12
 C:Genetics:
 A:Gene: CESP:F26F12.7
 A:Map position: 5
 A:introns: 110/3; 441/3; 801/2; 1244/3; 1693/2; 1784/1

Query Match 62.3%; Score 38; DB 2; Length 1829;
 Best Local Similarity 70.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISVPLTSPVC 10
 |||||
 DB 343 CIDPPLTEVP 352

RESULT 10
 B96814
 hypothetical protein T30F21.11 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96814
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A.
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewa-
 -ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mar-
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Ta-
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MUID:21016719
 A:Accession: B96814
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-184 <STO>
 A:Cross-references: GB:AE005173; NID:g4836878; PIDN:AAD30581.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T30F21.11
 A:Map position: 1

Query Match 60.7%; Score 37; DB 2; Length 184;
 Best Local Similarity 85.7%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 PLTSPVC 11
 |||||
 DB 150 PLTAVPC 156

RESULT 11
 H85135
 hypothetical protein AT4G12650 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H85135
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MUID:20083488
 A:Accession: H85135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-527 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267967; PIDN:CAB78308.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G12650
 A:Map position: 4

Query Match 60.7%; Score 37; DB 2; Length 527;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CISVPLT 7
 |||||
 DB 344 CISVPLT 350

RESULT 12
 F70592
 Probable IpgB protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: F70592
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gor-
 ; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holt

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: F70592
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-583 <COL>
 A:Cross-references: GB:295121; GB:AL123456; NID:g3261742; PIDN:CAB08345.1; PID:g314477;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: lpqB

Query Match 60.7%; Score 37; DB 2; Length 583;
 Best Local Similarity 70.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLTSPV 10
 |||||
 Db 16 CASVPTSTAP 25

RESULT 13
 C64707
 cation-transporting ATPase, P-type - *Helicobacter pylori* (strain 26695)
 C:Species: *Helicobacter pylori*
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 11-Jan-2000
 C:Accession: G64707
 R:Tomb, J.F.; White, O.; Kellavagge, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujili, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A:Reference number: A64520; MUID:97394467
 A:Accession: G64707
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-788 <TOM>
 A:Cross-references: GB:AE000648; GB:AE000511; NID:g2314670; PIDN:AAD08539.1; PID:g231467
 C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
 F:218-547/Domain: ATPase transduction domain homology <ATN>
 F:612-750/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 60.7%; Score 37; DB 2; Length 788;
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLTSPV 11
 |||||
 Db 456 C1SVPLTSPV 466

RESULT 14
 E71813
 probable component of cation transport for cbb3-type oxidase - *Helicobacter pylori* (stra
 C:Species: *Helicobacter pylori*
 A:Variety: strain j99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jan-2000
 C:Accession: E71813
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557
 A:Accession: E71813
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-788 <ARN>
 A:Cross-references: GB:AE001561; GB:AE001439; NID:g4156000; PIDN:AAD06962.1; PID:g415600

A:Experimental source: strain j99
 C:Genetics:
 A:Gene: fixI
 C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-bind
 F:612-750/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 60.7%; Score 37; DB 2; Length 788;
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLTSPV 11
 |||||
 Db 456 C1SVPLTSPV 466

RESULT 15
 T20160
 hypothetical protein T14G8.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T20160; T24924
 R:McMurray, A.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19231
 A:Accession: T20160
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1787 <WTL>
 A:Cross-references: EMBL:Z67881; PIDN:CAA91798.1; GSPDB:GN00028; CESP:T14G8.1
 A:Experimental source: clone C52G5
 R:Matthews, P.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19955
 A:Accession: T24924
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1787 <W12>
 A:Cross-references: EMBL:Z67884; PIDN:CAA91810.1; GSPDB:GN00028; CESP:T14G8.1
 C:Experimental source: clone T14G8
 C:Genetics:
 A:Gene: CESP:T14G8.1
 A:Map position: X
 A:Introns: 112/3; 453/3; 597/3; 815/2; 1258/3; 1682/2; 1709/3; 1764/1

Query Match 60.7%; Score 37; DB 2; Length 1787;
 Best Local Similarity 60.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLTSPV 10
 |||||
 Db 354 C1SVPLTSPV 363

Search completed: June 17, 2002, 15:58:21
 Job time: 190 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:26 ; Search time 44.84 Seconds
(without alignments)
9.499 Million cell updates/sec

Title: US-09-761-636A-7
Perfect score: 61
Sequence: 1 CISVPLTSPVC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	43	70.5	326	1 VEGD_RAT
2	43	70.5	354	1 VEGD_HUMAN
3	43	70.5	358	1 VEGD_MOUSE
4	39	63.9	207	1 TIM1_PIG
5	39	63.9	451	1 V110_FOXPV
6	39	63.9	1723	1 AIM1_HUMAN
7	38	62.3	503	1 YD39_YEAST
8	38	62.3	662	1 T952_MOUSE
9	38	62.3	663	1 T952_HUMAN
10	37	60.7	1787	1 CHD3_CAEEL
11	36	59.0	65	1 BBI2_SCHCO
12	36	59.0	530	1 SYK_MERIA
13	36	59.0	533	1 SYK_METAP
14	36	59.0	613	1 ADAS_TRYBB
15	35	57.4	207	1 TIM1_BOVIN
16	35	57.4	207	1 TIM1_SHEEP
17	35	57.4	363	1 ORC2_ARATH
18	35	57.4	756	1 K6PF_CAEEL
19	35	57.4	1041	1 TLR8_HUMAN
20	34	55.7	76	1 VEGD_BPMU
21	34	55.7	120	1 SY23_HUMAN
22	34	55.7	207	1 TIM1_HORSE
23	34	55.7	258	1 KPM1_ECOLI
24	34	55.7	258	1 KPM2_ECOLI
25	34	55.7	319	1 LDH_THESA
26	34	55.7	323	1 TKRA_ERMHE
27	34	55.7	349	1 YHHT_ECOLI
28	34	55.7	837	1 MYCCL_RAT
29	34	55.7	1097	1 EX5C_MYCTU
30	34	55.7	1233	1 VLI1_REOVY
31	34	55.7	1426	1 CUT2_MOUSE
32	34	55.7	1816	1 LMA4_HUMAN
33	34	55.7	2142	1 BAT2_HUMAN

34	34	55.7	2569	1 LMA3_MOUSE	Q6189 mus musculus
35	34	55.7	4753	1 LRP_CAEEL	Q0483 caenorhabditis
36	33.5	54.9	473	1 MDM2_XENLA	P56273 xenopus lae
37	33	54.1	100	1 PAHO_MOUSE	P10601 mus musculus
38	33	54.1	100	1 R14_NBPOL	Q9138 neproselmia
39	33	54.1	106	1 HOXL_AZOVI	P40392 azotobacter
40	33	54.1	108	1 HUPN_AZOC	P43955 azotobacter
41	33	54.1	205	1 TIM1_MOUSE	P12032 mus musculus
42	33	54.1	207	1 TIM1_HUMAN	P01033 homo sapien
43	33	54.1	207	1 TIM1_PAPCY	P49061 papio cynoc
44	33	54.1	218	1 TAL_RHLO	Q98970 rhizobium
45	33	54.1	257	1 DIS2_DICDI	P42530 dictyostell

ALIGNMENTS

RESULT	1	STANDARD;	PRT;	326 AA.
VEGD_RAT				
ID	035251:			
AC	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).			
DE	FIGF OR VEGFD.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
ON	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-Sprague-Dawley;			
RA	Yamada Y., Hirata Y., Nezu J., Shimane M.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor (By similarity).			
CC	-1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Secreted (By similarity).			
CC	-1- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS) bound by non-covalent interactions (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL: AF014827; AAB6557.1; -			
DR	HSSP: P15692; 1VP.			
DR	InterPro: IPR000072; PDGF.			
DR	Pfam: PF00341; PDGF. 1.			
DR	ProDom: PD001629; PDGF. 1.			
DR	SMART: SM00141; PDGF. 1.			
DR	PROSITE: PS00249; PDGF. 1.			
DR	PROSITE: PS50278; PDGF. 2; 1.			
KW	Mitogen; Growth factor; Glycoprotein; Signal; Repeat;			
KW	Cleavage on pair of basic residues; Multigene family.			
FT	SIGNAL	1	21	POTENTIAL.
FT	PROPEP	22	93	POTENTIAL.

```

FT CHAIN 94 210 VASCULAR ENDOTHELIAL GROWTH FACTOR D.
FT PROPEP 211 326 POTENTIAL.
FT DOMAIN 227 317 4 X 16 AA REPEATS OF C-X(10)-C-X-C-
  X(1,3)-C.
FT REPEAT 227 242 1 (APPROXIMATE).
FT REPEAT 263 278 2.
FT REPEAT 282 298 3.
FT REPEAT 306 317 4 (INCOMPLETE).
FT DISULFID 116 158 INTRACHAIN (BY SIMILARITY).
FT DISULFID 147 194 INTRACHAIN (BY SIMILARITY).
FT DISULFID 151 196 INTRACHAIN (BY SIMILARITY).
FT DISULFID 141 141 INTERCHAIN (BY SIMILARITY).
FT DISULFID 150 150 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 37112 MW; 1261AF373596C00 CRC64;

```

```

Query Match 70.5%; Score 43; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2 ISVPLTSVP 10
Db 173 ISVPLTSVP 181

```

```

RESULT 2
VEGD_HUMAN STANDARD; PRT: 354 AA.
ID VEGD_HUMAN STANDARD; PRT: 354 AA.
AC 043915;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
  growth factor) (Figf).
GN FIGF OR VEGFD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
  VEGF-D."
RL Genomics 42:483-488(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98140120; PubMed=9479493;
RA Rocchiolani M., Lestingi M., Luddi A., Orlandini M., Franco B.,
RT Rossi E., Balabio A., Zuffardi O., Oliviero S.;
RT "Human Figf: cloning, gene structure, and mapping to chromosome Xp22.1
  between the PIGA and the GRP genes."
RL Genomics 47:207-216(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118549; PubMed=9435229;
RA Achen M.G., Jeltsch M., Kukk E., Mäkinen T., Vitali A., Wilks A.F.,
RT Altalio K., Stacker S.A.;
RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
  tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
RP PROCESSING AND SEQUENCE OF 89-94; 100-105 AND 206-213.
RX MEDLINE=20011413; PubMed=10542248;
RA Stacker S.A., Stenvers K.L., Caesar C., Vitali A., Domagala T.,
RA Nice E.C., Roufail S., Simpson R.J., Moritz R., Karpanen T.,
RA Altalio K., Achen M.G.;

```

```

RT "Biosynthesis of vascular endothelial growth factor-D involves
  proteolytic processing which generates non-covalent homodimers."
RL J. Biol. Chem. 274:32127-32136(1999).
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
  and endothelial cell growth, stimulating their proliferation and
  migration and also has effects on the permeability of blood
  vessels. May function in the formation of the venous and lymphatic
  vascular systems during embryogenesis, and also in the maintenance
  of differentiated lymphatic endothelium in adults. Binds and
  activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
CC -1- TISSUE SPECIFICITY: Highly expressed in lung, heart, small
  intestine and fetal lung, and at lower levels in skeletal muscle,
  colon, and pancreas.
CC -1- PMW: Undergoes a complex proteolytic maturation which generates a
  variety of processed secreted forms with increased activity toward
  VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
  linked by disulfide bonds before secretion. The fully processed
  VEGF-D is composed mostly of two VEGF homology domains (VHDS)
  bound by non-covalent interactions.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D89630; BA24264.1; -
DR EMBL: Y12863; CAA73370.1; -
DR EMBL: Y12864; CAA73371.1; -
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12866; CAA73371.1; JOINED.
DR EMBL: Y12867; CAA73371.1; JOINED.
DR EMBL: Y12868; CAA73371.1; JOINED.
DR EMBL: Y12869; CAA73371.1; JOINED.
DR EMBL: Y12870; CAA73371.1; JOINED.
DR EMBL: A1000185; CAA03942.1; -
DR KIM: 300091; -
DR HSSP: P15692; IVP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF. 1.
DR ProDom: PD001629; PDGF. 1.
DR SMART: SM00141; PDGF. 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 88
FT CHAIN 89 205
FT PROPEP 206 354
FT DOMAIN 222 318
FT REPEAT 222 237
FT REPEAT 258 273
FT REPEAT 277 293
FT REPEAT 301 318
FT DISULFID 111 153
FT DISULFID 142 189
FT DISULFID 146 191
FT DISULFID 136 136
FT DISULFID 145 145
FT CARBOHYD 155 155
FT CARBOHYD 185 185
FT CARBOHYD 287 287
SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

```

Query Match

70.5%; Score 43; DB 1; Length 354;

Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
|||||
Db 168 ISVPLTSVP 176

RESULT 3
VEG-D_MOUSE STANDARD; PRT: 358 AA.
AC P97946;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (Figf).
DE FIGF OR VEGFD.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=FltProblast;
RX MEDLINE=97030254; PubMed=8876195;
RA Oriandini M., Marconini L., Ferruzzi R., Oliviero S.;
RT Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor, VEGF-D.";
RL Genomics 42:483-488(1997).
[3]
RP DEVELOPMENTAL STAGE.
RX MEDLINE=98288130; PubMed=9622638;
RA Avastgiato V., Oriandini M., Acampora D., Oliviero S., Simeone A.;
RT "Embryonic expression pattern of the murine figf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family.";
RL Mech. Dev. 73:221-224(1998).
[4]
RP RECEPTOR SPECIFICITY.
RX MEDLINE=21276411; PubMed=11279005;
RA Baldwin M.E., Catmel B., Nice E.C., Roufall S., Hall N.E., Stenvers K.L., Karkkainen M.J., Altalo K., Stacker S.A., Achen M.G.;
RT "The specificity of receptor binding by vascular endothelial growth factor-d is different in mouse and man.";
RL J. Biol. Chem. 276:19166-19171(2001).
-1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor.
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
CC -1- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.
CC -1- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several body structures and organs of the embryo such as limb buds, acoustic ganglion, teeth, heart, anterior pituitary as well as lung and kidney mesenchyme, liver, derma, and peritoneum of the vertebral column.
CC -1- INDUCTION: By the transcription factor c-fos.
CC -1- PTH: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward

CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS) bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC -----
CC EMBL: X99572; CAA67892.1; -
CC EMBL: D89628; BAA14002.1; -
CC HSSP: P15692; 1VPP.
CC MGD: MGI:108037; Figf.
CC InterPro: IPR000072; PDGF.
CC Pfam: PF00341; PDGF; 1.
CC ProDom: PD001629; PDGF; 1.
CC SMART: SM00141; PDGF; 1.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS50278; PDGF_2; 1.
CC Mitogen; Growth factor; Glycoprotein; Signal; Repeat; Cleavage on pair of basic residues; Multigene family.
KW SIGNAL 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 358
FT DOMAIN 227 323
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 306 323
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 358 AA; 40908 MW; 6636B17F5F07037C CRC64;

Query Match 70.5%; Score 43; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
|||||
Db 173 ISVPLTSVP 181

RESULT 4
TIM1_PIG STANDARD; PRT: 207 AA.
AC P35624; Q9T83; Q9T89;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metalloprotease inhibitor 1 precursor (TIMP-1).
GN TIMP1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=92201478; PubMed=1312961;

```

DT 01-JUN-1994 (Rel. 29, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein FFPV110.
GN FFPV110.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxId=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C. L., Tjulan E. R., Lu Z., Zsak L., Kutish G. F., Rock D. L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
RN [2]
RP SEQUENCE OF 1-151 FROM N.A.
RX MEDLINE=93139784; PubMed=8380837;
RA Ogawa R., Calvert J. G., Yanagida N., Nazerian K. ;
RT "Insertional inactivation of a fowlpox virus homologue of the
RL vaccinia virus F12L gene inhibits the release of enveloped virions."
CC J. Gen. Virol. 74:55-64(1993).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES F1L FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF1981100; AAA4454.1; -
DR M88588; AAA47188.1; -.
DR PIR: S27935; S27935.
DR PIR: P00506; P00506.
QO SEQUENCE 451 AA; 52010 MW; 664623EA83DA331 CRC64;

```

```

Query Match          63.9%; Score 39; DB 1; Length 451;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY      1 C1SVPLTSSVP 10
      ||::||::||:|
Db      90 C1NPIDSIP 99

RESULT 6
A1M1_HUMAN
AC      A1M1_HUMAN          STANDARD;          PRT; 1723 AA.
Q9YAK1; O00296;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Absent in melanoma 1 protein.
GN      A1M1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
[1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Liver;
RX      MEDLINE=97250519; PubMed=9096375;
RA      Ray M.E., Wistow G., Su Y.A., Meltzer P.S., Trent J.M.;
RT      "A1M1, a novel non-lens member of the betagamma-crystallin
RT      superfamily, is associated with the control of tumorigenicity in human
RT      malignant melanoma.";
RL      Proc. Natl. Acad. Sci. U.S.A. 94:3229-3234(1997).
CC      -1- FUNCTION: MAY FUNCTION AS SUPPRESSOR OF MALIGNANT MELANOMA. IT MAY
CC      EXERT ITS EFFECTS THROUGH INTERACTIONS WITH THE CYTOSKELETON.
CC      -1- SIMILARITY: BELONGS TO THE BETA/GAMMA-CRYSTALLIN FAMILY. CONTAINS

```


DR PROSITE: PS00598; CHROMO_1; FALSE_NEG.
 DR PROSITE: PS50013; CHROMO_2; 1.
 KW Hypothetical protein; Nuclear protein; Repeat; Helicase; DNA-binding;
 FT ZN_BINDING; Zinc-finger.
 FT ZN_BINDING 59 62 POLY-LYS.
 FT ZN_BINDING 268 309 PHD-TYPE 1.
 FT ZN_BINDING 331 372 PHD-TYPE 2.
 FT ZN_BINDING 373 476 CHROMO_1.
 FT DOMAIN 501 583 CHROMO_2.
 FT DOMAIN 1287 1291 POLY-ARG.
 FT NE_BIND 641 648 ATP (POTENTIAL).
 FT SITE 763 766 DEAH BOX.
 SQ SEQUENCE 1787 AA; 205254 MW; 1EFCF1FECE59740 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 1787;
 Best Local Similarity 60.0%; Pred. No. 97;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISVPLTSPV 10
 11 111 :1
 DB 354 CIDPELTSP 363

RESULT 11
 BL12_SCHCO STANDARD: PRT: 65 AA.
 ID BL12_SCHCO
 AC P78743;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Mating-type pheromone BBP1(2) precursor.
 GN BBP1(2).
 OS Schizophyllum commune (Bracket fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Stereales; Schizophyllaceae; Schizophyllum.
 OC NCBI_TaxID=5334;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4-40;
 RA MEDLINE=97321284; PubMed=9178005;
 RA Vallancourt L.J., Raudaskoski M., Specht C.A., Raper C.A.;
 RT "Multiple genes encoding pheromones and a pheromone receptor define
 the B beta 1 mating-type specificity in Schizophyllum commune.";
 RL Genetics 146:541-551(1997).
 CC -1- FUNCTION: ACTIVATES B-REGULATED DEVELOPMENT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U74495; AAB41860.1; -
 DR prenylation; Lipoprotein; Pheromone.
 KW PROPEP 1 2 POTENTIAL.
 FT PEPTIDE 2 62 MATING-TYPE PHEROMONE BBP1(2).
 FT PROPEP 63 65 REMOVED IN MATURE FORM (POTENTIAL).
 FT LIPID 62 62 FARNESYL (POTENTIAL).
 SQ SEQUENCE 65 AA; 6780 MW; 3341013317AD6065 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 65;
 Best Local Similarity 66.7%; Pred. No. 5;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 SVPLTSPV 11
 1111 :1
 DB 42 SAPLEAPC 50

RESULT 12
 SYK_METUA STANDARD: PRT: 530 AA.
 ID SYK_METUA
 AC O57959;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LYSRS).
 GN LYS OR M0539.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OC NCBI_TaxID=2190;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RA MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Keriavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Hansen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
 CC + L-lysyl-tRNA(Lys).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67503; AAB98532.1; -
 DR TIGR: M0539;
 DR InterPro: IPR001412; tRNA-synt-1.
 DR InterPro: IPR002904; tRNA-synt-lys_1.
 DR Pfam: PF01921; tRNA-synt_1f; 1.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 28 36 "HIGH" REGION.
 FT SITE 278 282 "KMSKS" REGION.
 SQ SEQUENCE 530 AA; 61921 MW; 5811837C8A349E9C CRC64;

Query Match 59.0%; Score 36; DB 1; Length 530;
 Best Local Similarity 50.0%; Pred. No. 42;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ISVPLTSPV 11
 1111 :1
 DB 84 IGMPLEIPC 93

RESULT 13
 SYK_METMP STANDARD: PRT: 533 AA.
 ID SYK_METMP
 AC O30522;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine--tRNA ligase) (LYSRS).
 GN LYS.
 OS Methanococcus maripaludis.

CC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus;
OX NCBI_TaxID=39152;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22.
RC STRAIN-JJ;
RX MEDLINE=98016282; PubMed=9353192;
RA Ibbas M., Morgan S., Curnow A.W., Pridmore D.R., Volhnecht U.C.,
RT Gardner W., Lin W., Moese C.R., Soell D.;
RT "A euryarchaeal Lysyl-tRNA synthetase: resemblance to class I
RT synthetases.";
RT Science 278:1119-1122(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-lysine + tRNA(Lys) = AMP + diphosphate
CC + L-lysyl-tRNA(Lys).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF009824; AAB87410.1; -
DR InterPro: IPR001412; tRNA-syn-I.
DR InterPro: IPR002904; tRNA-syn-Lys-1.
DR Pfam: PF01921; tRNA-syn-Lf; 1.
DR PROSITE: PS00178; AA-TRNA-LIGASE.1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 28 36 "HIGH" REGION.
FT SITE 278 282 "KMSK" REGION.
FT COMPACT 11 11 I -> L (IN REF. 1; AA SEQUENCE).
FT SEQUENCE 533 AA; 61273 MW; 007FA868A57A0AC2 CRC64;
QY 2 ISVPLTSVPC 11
Db 84 IGMPLESDIC 93
RESULT 14
ADAS-TRYBB STANDARD; PRT; 613 AA.
ID ADAS-TRYBB
DC 097157;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alkylidihydroxyacetonephosphate synthase (EC 2.5.1.26) (Alkyl-DHAP
DE synthase) (Alkylglycerone-phosphate synthase).
OS Trypanosoma brucei brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN (1)
RP SEQUENCE FROM N.A.
RA Zomer A.M.W., Michels P.A.M., Opperdoes F.R.;
RT "Molecular characterization of Trypanosoma brucei alkyl-
RT dihydroxyacetonephosphate synthase.";
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 1-acyl-glycerone 3-phosphate + a long-chain
CC alcohol = 1-alkyl-glycerone 3-phosphate + a long-chain acid anion.
CC -1- COFACTOR: FAD (By similarity).
CC -1- PATHWAY: ETHER LIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FAD-BINDING OXIDOREDUCTASE/TRANSFERASE
CC FAMILY 4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF119091; AAD19697.1; -
DR InterPro: IPR004113; FAD-oxidase-C.
DR InterPro: IPR001575; Oxid_FAD_bind.
DR Pfam: PF02913; FAD-oxidase-C; 1.
DR Pfam: PF01565; FAD-binding_4; 1.
KW Lipid synthesis; Transferase; Flavo-protein; FAD; Peroxisome.
FT SITE 611 613 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT SEQUENCE 613 AA; 69067 MW; 27BA9FBF2EB94E3A CRC64;
QY 2 ISVPLTSVPC 11
Db 270 VETPLTSRPC 279
RESULT 15
TIML_BOVIN STANDARD; PRT; 207 AA.
ID TIML_BOVIN
AC P20414; Q9TVB0;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metalloprotease inhibitor 1 precursor (TIMP-1) (Embryogenin-1) (EG-
DE 1).
OS Bos taurus (Bovine).
GN TIMP1.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=90365711; PubMed=2393392;
RA Freudenstein J., Wagner S., Luck R.M., Einspanier R., Scheit K.H.;
RT "mRNA of bovine tissue inhibitor of metalloproteinase: sequence and
RT expression in bovine ovarian tissue.";
RT Biochem. Biophys. Res. Commun. 171:250-256(1990).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=94257757; PubMed=8199264;
RA Satoh T., Kobayashi K., Yamashita S., Kikuchi M., Sendai Y., Hoshi H.;
RT "Tissue inhibitor of metalloproteinases (TIMP-1) produced by
RT granulosa and oviduct cells enhances in vitro development of bovine
RT embryo.";
RL Biol. Reprod. 50:835-844(1994).
RN (3)
RP SEQUENCE OF 25-191 FROM N.A.
RC TISSUE-SKELETAL muscle;
RA Balcerzak D., Quenengesser L., Dixon W.T., Baracos V.E.;
RT "Involvement of fibroblasts and muscle cells in the expression of an
RT extracellular proteolytic cascade in bovine skeletal muscle.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN (4)
RP PRELIMINARY SEQUENCE OF 24-69.
RX MEDLINE=90008914; PubMed=2551903;
RA de Clerck Y.A., Veau T.D., Ratzkin B.J., Lu H.S., Langley K.E.;
RT "Purification and characterization of two related but distinct
RT metalloproteinase inhibitors secreted by bovine aortic endothelial
RT cells.";
RL J. Biol. Chem. 264:17445-17453(1989).
CC -1- FUNCTION: COMPLEXES WITH METALLOPROTEINASES (SUCH AS COLLAGENASES)
CC AND IRREVERSIBLY INACTIVATES THEM.

```

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: THE ACTIVITY OF TIMP-1 IS DEPENDENT ON THE PRESENCE OF
CC DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE TIMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60073; AAA30784.1; -
DR EMBL; S70841; AAB30892.1; -
DR EMBL; AF144763; AAD30303.1; -
DR PIR; A35685; A35685
DR PIR; B34468; B34468
DR HSSP; P01033; IUEA
DR InterPro; IPR001820; TIMP
DR Pfam; PF00965; TIMP; 1
DR SMART; SM00206; TIMP; 1
DR PROSITE; PS00288; TIMP; 1
KW Glycoprotein; Metalloprotease inhibitor; Erythrocyte maturation;
KW Signal.
FT CHAIN 1 23 METALLOPROTEINASE INHIBITOR 1.
FT DISULFID 24 207 BY SIMILARITY.
FT DISULFID 24 93 BY SIMILARITY.
FT DISULFID 26 122 BY SIMILARITY.
FT DISULFID 36 147 BY SIMILARITY.
FT DISULFID 150 197 BY SIMILARITY.
FT DISULFID 155 160 BY SIMILARITY.
FT DISULFID 168 189 BY SIMILARITY.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 207 AA; 23031 MW; E672BEE2E865F3F7 CRC64;

```

```

Query Match 57.4%; Score 35; DB 1; Length 207;
Best Local Similarity 45.5%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 CISVPLTSVPC 11
Db 150 CTVFPCCSIPC 160

```

Search completed: June 17, 2002, 16:24:27
Job time: 1556 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:35 ; Search time 172.85 Seconds

(without alignments)
11.009 Million cell updates/sec

Title: US-09-761-636a-7
Perfect score: 61
Sequence: 1 CISVPLTSVPC 11

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	70.5	122	6	09GLX1	09glx1 bos taurus
2	43	70.5	326	11	09LZE4	09lze4 rattus norv
3	42	68.9	348	5	09GUB7	09gub7 caenorhabd
4	40	65.6	208	5	09NF93	09nfv3 leishmania
5	40	63.9	169	11	09Z1H6	09z1h6 mus musculu
6	39	63.9	169	17	09Y9B5	09y9b5 aeropyrum p
7	39	63.9	318	12	09PY69	09py69 soybean mos
8	39	63.9	661	10	09XG05	09xg05 oryza sativ
9	39	63.9	1062	4	09NSW2	09nsw2 homo sapien
10	39	63.9	1411	4	09GQW6	09gqw6 homo sapien
11	39	63.9	1513	3	09HDV4	09hdv4 schizosacch
12	38	62.3	207	3	009411	009411 arthropoda
13	38	62.3	318	12	068691	068691 hepatitis c
14	38	62.3	324	5	016608	016608 caenorhabd
15	38	62.3	616	5	020852	020852 caenorhabd
16	38	62.3	638	10	09FY08	09fy08 arabidopsis

17	38	62.3	1829	5	019815	019815 caenorhabd
18	37	60.7	156	11	09D664	09d664 mus musculu
19	37	60.7	178	2	09L5F5	09l5f5 salmonella
20	37	60.7	184	10	09SYM6	09sym6 arabidopsis
21	37	60.7	251	12	09Q913	09q913 Shope fibro
22	37	60.7	251	12	09Q8N8	09q8n8 myxoma viru
23	37	60.7	401	10	093YR2	093yr2 arabidopsis
24	37	60.7	527	10	09SU21	09su21 arabidopsis
25	37	60.7	583	16	005889	005889 mycobacteri
26	37	60.7	788	16	026033	026033 helicobacte
27	37	60.7	788	16	09ZJB7	09zjb7 helicobacte
28	37	60.7	822	5	09VL64	09vl64 drosophila
29	37	60.7	887	5	016102	016102 drosophila
30	37	60.7	1060	4	09P1Z7	09p1z7 homo sapien
31	37	60.7	1329	11	091ZV8	091zv8 mus musculu
32	37	60.7	1331	4	096PE1	096pe1 homo sapien
33	37	60.7	1838	5	09VMJ7	09vmj7 drosophila
34	37	60.7	2327	13	09IBG7	09ibg7 xenopus lae
35	36	59.0	165	17	09YF66	09yfr6 aeropyrum p
36	36	59.0	174	2	065940	065940 escherichia
37	36	59.0	198	11	09CSA8	09csa8 mus musculu
38	36	59.0	199	11	09COR2	09cor2 mus musculu
39	36	59.0	245	4	09NVP6	09nvp6 homo sapien
40	36	59.0	261	5	020542	020542 caenorhabd
41	36	59.0	271	11	09D4L6	09d4l6 mus musculu
42	36	59.0	334	13	090468	090468 brachydanio
43	36	59.0	337	4	09P296	09p296 homo sapien
44	36	59.0	342	3	013319	013319 metarhizium
45	36	59.0	397	5	076862	076862 drosophila

ALIGNMENTS

RESULT 1
ID 09GLX1 PRELIMINARY: PRT: 122 AA.
AC 09GLX1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR-D (FRAGMENT).
GN VEGF-D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP TISSUE=LUNG;
RC SEQUENCE FROM N.A.
RA Mandriota S.J., Pepper M.S.;
RU Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099135; AAG29747.1; -;
DR InterPro; IPR004153; CXXC.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF03128; CXXC; 2.
DR ProDom; PD001629; PDGF; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 122 AA: 13820 MW; CC504B00E29D54EB CRC64;

Query Match 70.5%; Score 43; DB 6; Length 122;
Best Local Similarity 100.0%; Pred No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISVPLTSVP 10
|||||
DB 1 ISVPLTSVP 9

RESULT 2

```

Q91ZE4
ID Q91ZE4 PRELIMINARY; PRT; 326 AA.
AC Q91ZE4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VEGF-D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Kirkin V., Mazitschek R., Krishnan J., Steffen A., Waltenberger J.,
RA Pepper M.S., Giannis A., Sleeman J.P.;
RT "Characterization of indolinones which specifically inhibit VEGF-C and
RT VEGF-D-induced activation of VEGFR-3 but not VEGFR-2."
RT Eur. J. Biochem. 0:0-0(2001).
DR EMBL: AY032728; AAK96008.1; -.
SQ SEQUENCE 326 AA; 37106 MW; D7CAEBA6C9FAB7D CRC64;

Query Match
Best Local Similarity 100.0%; Score 43; DB 11; Length 326;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISVPLTSVP 10
DB 173 ISVPLTSVP 181

RESULT 3
ID Q9GUB7 PRELIMINARY; PRT; 348 AA.
AC Q9GUB7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 39.9 KDA PROTEIN.
GN F25E5.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wilson R., Bradshaw H.;
RT "The sequence of C. elegans cosmid F25E5."
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF078157; AAG24081.1; -.
DR InterPro: IPR003003; 7TM_chemorecept_2.
DR Pfam: PF01604; 7tm_5; 1.
DR Hypothetical protein.
SQ SEQUENCE 348 AA; 39882 MW; 456DEC8833CF3176 CRC64;

```

```

Query Match
Best Local Similarity 68.9%; Score 42; DB 5; Length 348;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CISVPLTSVPC 11
DB 33 CISVPLTSVPC 43

RESULT 4
ID Q9NF93 PRELIMINARY; PRT; 208 AA.
AC Q9NF93;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL TRANSMEMBRANE PROTEIN L8032.05A.
GN L8032.05A.
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RX MEDLINE-98146435; PubMed-9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the leishmania major Friedlin genome."
RT Genome Res. 8:135-145(1998).
DR EMBL: AL139794; CAC22646.1; -.
KW Transmembrane.
SQ SEQUENCE 208 AA; 23200 MW; 11DB7F6DDDC566A3 CRC64;

Query Match
Best Local Similarity 65.6%; Score 40; DB 5; Length 208;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CISVPLTSVPC 11
DB 96 CISVPLTSVPC 106

RESULT 5
ID Q9Z1H6 PRELIMINARY; PRT; 782 AA.
AC Q9Z1H6;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NUCLEAR PROTEIN NP95 (NUCLEAR ZINC FINGER PROTEIN NP95).
GN NP95 OR UHRF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PRE-TCELL;
RX MEDLINE-9909250; PubMed-9880673;
RA Fujimori A., Matsuda Y., Takemoto Y., Hashimoto Y., Kubo E., Araki R.,
RA Fukumura R., Mita K., Tatsumi K., Muto M.;
RT "Cloning and mapping of Np95 gene which encodes a novel nuclear
RT protein associated with cell proliferation."
RT Mamm. Genome 9:1032-1035(1998).
RN [2]
RP SEQUENCE FROM N.A.

```


RA Davaport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RT "LMO2-induced T cell leukemias overexpress Np95, a gene containing
RING and PHD zinc fingers and an ubiquitin-like domain."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: D87908; CAB20641.1; -
DR EMBL: AF274046; AAK55743.1; -
DR HSSP: Q15843; 1MDJ.
DR MGD: MGI:1338889; Np95.
DR InterPro: IPR003105; G9a.
DR InterPro: IPR000566; Lipoaln_cytfabp.
DR InterPro: IPR001965; PHD.
DR InterPro: IPR000626; Ubiquitin.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF02182; G9a; 1.
DR Pfam: PF00628; PHD; 1.
DR Pfam: PF00240; ubiquitin; 1.
DR PRINTS: PR00348; UBQUITIN.
DR SMART: SM00466; G9a; 1.
DR SMART: SM00249; PHD; 1.
DR SMART: SM00184; RING; 2.
DR SMART: SM00213; UBO; 1.
DR PROSITE: PS00213; LIPOALIN; UNKNOWN_1.
DR PROSITE: PS00518; ZINC_FINGER_C3HC4; UNKNOWN_1.
DR PROSITE: 782 AA; 88303 MW; DCSEDFCDF69619B CRC64;
SQ

Query Match 65.6%; Score 40; DB 11; Length 782;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CISVPLTSVP 10
|: |||||
Db 349 CLKPLTSVP 358

RESULT 6
ID Q9Y9B5 PRELIMINARY; PRT; 169 AA.
AC Q9Y9B5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 17.4 KDA PROTEIN APE2372.
GN APE2372.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RE MEDLINE=99310339; PubMed=10382866;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kuoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000064; BAA81385.1; -
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 169 AA; 17432 MW; 9BDB7147C7B1426A CRC64;

Query Match 63.9%; Score 39; DB 17; Length 169;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CISVPLTSVP 10
|: ||: |||

Db 57 CIGIPVASVP 66
RESULT 7
ID Q9PY69 PRELIMINARY; PRT; 318 AA.
AC Q9PY69;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P1 PROTEASE (FRAGMENT).
OS Soybean mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Polyvirus.
OX NCBI_TaxID=12222;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHI;
RA Latorre I.J., Dowler L.L., Hartman G.L.;
RT "Natural variability of soybean mosaic virus."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF200544; AAF13954.1; -
DR MEROPS: S30.UNW; -
DR InterPro: IPR002540; Poly_P1.
DR Pfam: PF01577; Poly_P1; 1.
KW Protease.
FT NON_TER 1 1
FT NON_TER 318 318
SQ SEQUENCE 318 AA; 36291 MW; 0D2C1B46035AEAF1 CRC64;

Query Match 63.9%; Score 39; DB 12; Length 318;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 ISVPLTSVPC 11
|||||: |||
Db 20 ISVPLTHIPC 29

RESULT 8
ID Q9XG05 PRELIMINARY; PRT; 661 AA.
AC Q9XG05;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ESTS A0064813(BA0579).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaristidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC
RT clone: p0026F07."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP000364; BAA81763.1; -
DR InterPro: IPR004240; EMP70.
DR Pfam: PF02990; EMP70; 1.
SQ SEQUENCE 661 AA; 74533 MW; 612242C132F68B8E CRC64;

Query Match 63.9%; Score 39; DB 10; Length 661;
Best Local Similarity 88.9%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CISVPLTSV 9
|||||||: |
Db 478 CISVPLTLV 486

RESULT 9
Q9NSM2 PRELIMINARY; PRT; 1062 AA.
AC Q9NSM2
DT 01-OCT-2000 (TREMBLER, 15, Created)
DT 01-OCT-2000 (TREMBLER, 15, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE HYPOTHEICAL 118.8 KDA PROTEIN.
GN DKE2P434N1427.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Bloecher H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RL Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137701; CAB70882.1; -
DR HSSP; P02649; INFO.
DR InterPro; IPR000719; Euk_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
KW ATP-binding; Hypothetical protein; Transferase.
SQ SEQUENCE 1062 AA; 118780 MW; A4B2B359EEC9CAAD CRC64;

Query Match 63.9%; Score 39; DB 4; Length 1062;
Best Local Similarity 66.7%; Pred. NO. 75;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLTSPVC 11
Db 25 SLPSTSPC 33

RESULT 10
Q96QW6 PRELIMINARY; PRT; 1411 AA.
AC Q96QW6
DT 01-DEC-2001 (TREMBLER, 19, Created)
DT 01-DEC-2001 (TREMBLER, 19, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE DJ448K1.1.1 (ABSENT IN MELANOMA 1, ISOFORM 1) (FRAGMENT).
GN ATM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Cobley V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359292; CAC44022.1; -
FT NON_TER 1411 1411
SQ SEQUENCE 1411 AA; 154110 MW; C1B36DD23F428C67 CRC64;

Query Match 63.9%; Score 39; DB 4; Length 1411;
Best Local Similarity 45.5%; Pred. NO. 98;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CISVPLTSPVC 11
Db 395 CVOSPLTSPVC 405

RESULT 11
Q9HDV4 PRELIMINARY; PRT; 1513 AA.

AC Q9HDV4;
DT 01-MAR-2001 (TREMBLER, 16, Created)
DT 01-MAR-2001 (TREMBLER, 16, Last sequence update)
DT 01-JUN-2001 (TREMBLER, 17, Last annotation update)
DE PUTATIVE TRANSCRIPTIONAL REGULATORY PROTEIN, PHD FINGER.
GN SPBP19A11.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Aert R., Robben J., Wellens I., Grymonprez B., Volckaert G., Wood V.,
RL Rajandream M.A., Barrell B.G., et al.
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL512495; CAC19756.1; -
DR InterPro; IPR001606; ARID.
DR InterPro; IPR003345; CytC_heme_bind.
DR InterPro; IPR001965; PHD.
DR InterPro; IPR001841; ZnF_Cing.
DR Pfam; PF02373; jmjC; 1.
DR Pfam; PF00628; PHD; 3.
DR SMART; SM00501; BRIGH1.
DR SMART; SM00249; PHD; 3.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 1513 AA; 172257 MW; B8787624F9E7DFAD CRC64;

Query Match 63.9%; Score 39; DB 3; Length 1513;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CISVPLTSPVC 10
Db 297 CLDPPLTSPIC 306

RESULT 12
Q09411 PRELIMINARY; PRT; 207 AA.
AC Q09411
DT 01-JUL-1997 (TREMBLER, 04, Created)
DT 01-JAN-1998 (TREMBLER, 05, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE CHITIN SYNTHASE 1 (FRAGMENT).
OS Arthropoda; Insecta; Diptera; Muscidae; Tephritidae; Anthomyiidae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Arthrodermataceae; Arthroderma.
OX NCBI_TaxID=63401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VUT-4006;
RX MEDLINE=98131180; PubMed=9470404;
RA Kano R., Nakamura Y., Watari T., Watanabe S., Takahashi H.,
RT Tsujimoto H., Hasegawa A.;
RT "Phylogenetic analysis of 8 dermatophyte species using chitin synthase
gene 1 sequences."
RL Mucosae 40:411-414(1997).
DR EMBL; AB003559; BAA20947.1; -
DR InterPro; IPR002923; Chitin_synth.
DR Pfam; PF01644; Chitin_synth; 1.
DR ProDom; PD002998; Chitin_synth; 1.
FT NON_TER 207 207
SQ SEQUENCE 207 AA; 22580 MW; 47B2E8FBD6499541 CRC64;

Query Match 62.3%; Score 38; DB 3; Length 207;
Best Local Similarity 87.5%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 VPLTSPVC 11
11111111
Db 166 VPLTSSPC 173

RESULT 13

068691 PRELIMINARY; PRT; 318 AA.
ID 068691;
AC 068691;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CORE PROTEIN (FRAGMENT).
OS Hepatitis C virus type 2c.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31651;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96118171; PubMed=8578855;
RA Stuyver L., Wyseur A., van Arnhem W., Lunel F., Laurent-Pulig P.,
RA Pawlotsky J.-M., Kleter B., Bassit L., Nkengasong J., van Doorn L.-J.,
RA Maertens G.,
RT "Hepatitis C virus genotyping by means of 5'-UTR/core line probe assays
RT and molecular analysis of untypeable samples.";
RL Virus Res. 38:137-157(1995).
DR EMBL; L38324; AAC42179.1;
FT CHAIN 1 >191 CORE PROTEIN.
FT NON_TER 192 >318 EL PROTEIN.
FT SEQUENCE 318 AA; 34463 MW; E7A5F7DE56A34FC CRC64;
SQ

Query Match 62.3%; Score 38; DB 12; Length 318;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVPLTSPV 9
1111111111
Db 184 C1SVPLTSPV 192

RESULT 14

016608 PRELIMINARY; PRT; 324 AA.
ID 016608;
AC 016608;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 37.0 KDA PROTEIN.
GN C49D10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Henkhaus J., Wohldmann P., Beck C.,
RT "The sequence of C. elegans cosmid C49D10.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016665; AAC71178.1;
DR InterPro; IPR003003; 7TM_chemorecept_2.
DR InterPro; IPR000168; 7TM_nematode.
DR Pfam; PF01604; 7tm_5; 1.
KW Hypothetical protein.
SQ SEQUENCE 324 AA; 37015 MW; AA9BB3BD4D2DD926 CRC64;

Query Match 62.3%; Score 38; DB 5; Length 324;
Best Local Similarity 54.5%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 C1SVPLTSPVC 11
1111111111
Db 70 C1GVPLTSPVC 80

RESULT 15

020852 PRELIMINARY; PRT; 616 AA.
ID 020852;
AC 020852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 66.0 KDA PROTEIN.
GN F55G1.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Murray J., Le T.T.;
RT "The sequence of C. elegans cosmid F55G1.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58750; AAB00653.1;
DR HSSP; P01132; 1EGF.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00008; EGF_8.
DR SMART; SM000179; EGF_CA_1.
DR SMART; SM00001; EGF-like; 7.
DR PROSITE; PS00010; ASX_HYDROXYL_1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_7.
DR PROSITE; PS01186; EGF_2; 2.
DR EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
SQ SEQUENCE 616 AA; 66017 MW; 06B8DA94F055C6ED CRC64;

Query Match 62.3%; Score 38; DB 5; Length 616;
Best Local Similarity 54.5%; Pred. No. 69;

	Matches	6;	Conservative	1;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	1	CISVPLTSVPC	11							
		1:	1111							
Db	362	CVGNPCTSEPC	372							

Search completed: June 17, 2002, 16:23:37
Job time: 1611 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 17, 2002, 16:02:12 ; Search time 227.35 Seconds
(Without alignments)
5.374 Million cell updates/sec

Title: US-09-761-636a-7
Perfect score: 61
Sequence: 1 CISVPLRVSVC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	100.0	11	22	AAU04526	VEGF based monocyc
2	59	96.7	11	22	AAU04542	VEGF based monocyc
3	57	93.4	11	22	AAU04545	VEGF based monocyc
4	54	88.5	11	22	AAU04543	VEGF based monocyc
5	52	85.2	11	22	AAU04544	VEGF based monocyc
6	45.5	74.6	10	22	AAU04532	VEGF based monocyc
7	43	70.5	96	22	AAU04520	Human VEGF-D amino
8	43	70.5	109	20	AAU23889	Human vascular end
9	43	70.5	109	21	AAU11931	Human truncated VE
10	43	70.5	178	20	AAU08287	Human growth facto
11	43	70.5	321	19	AAU53243	Mus musculus vascu

12	43	70.5	325	19	AAU53240	Homo sapiens vascu
13	43	70.5	325	22	AAU97572	Human VEGF-D prote
14	43	70.5	326	19	AAU44296	Rat vascular endot
15	43	70.5	337	20	AAU08286	Human growth facto
16	43	70.5	354	19	AAU49036	Human zvegf2 growt
17	43	70.5	354	19	AAU53241	Homo sapiens vascu
18	43	70.5	354	19	AAU44293	Human VEGD protein
19	43	70.5	354	21	AAU10649	Human VEGF-D prote
20	43	70.5	354	21	AAU29049	Human prepro-Vascu
21	43	70.5	354	21	AAU70750	Human vascular end
22	43	70.5	354	21	AAU70983	Polypeptide for hu
23	43	70.5	354	22	AAU08441	Human vascular end
24	43	70.5	354	22	AAU70685	Human VEGF-D1 prot
25	43	70.5	354	22	AAU97573	Human VEGF-D. Hom
26	43	70.5	354	22	AAU37606	Murine c-Fos Induc
27	43	70.5	358	18	AAU14992	Mus musculus vascu
28	43	70.5	358	19	AAU53242	Mus vascular end
29	43	70.5	358	19	AAU44295	Human c-Fos induce
30	43	70.5	620	18	AAU14994	Human secreted pro
31	43	70.5	39	20	AAU70909	VEGF based monocyc
32	41	67.2	9	22	AAU04533	Novel human secret
33	41	67.2	60	22	AAU33238	Human secreted pro
34	41	67.2	121	21	AAU04076	Human signal pepti
35	40	65.6	80	21	AAU87243	Human pro polypept
36	40	65.6	80	22	AAU29311	Human polypeptide
37	39	63.9	54	22	AAU06993	Proionibacterium
38	39	63.9	440	22	AAU40825	Human bone marrow
39	39	63.9	1062	21	AAU40294	Human protein Ssq
40	39	63.9	1637	22	AAU00916	Human polypeptide
41	39	63.9	2263	22	AAU79000	Human polypeptide
42	38	62.3	67	22	AAU56908	Human polypeptide
43	38	62.3	117	22	AAU010752	Human polypeptide
44	38	62.3	193	22	ABG11976	Human polypeptide
45	38	62.3	207	16	AAU65004	Pig tissue inhibit

ALIGNMENTS

RESULT 1

AAU04526 standard; Peptide: 11 AA.

AAU04526:

26-SEP-2001 (first entry)

VEGF based monocyclic peptide 3.

Human: VEGF: vascular endothelial growth factor; angiogenesis;
neovascularisation; lymphangiogenesis; psoriasis; tumour;
diabetes induced neovascular sequelae; rheumatoid arthritis;
diabetic retinopathy; chronic inflammation; cyclic.

Synthetic.

Key Location/Qualifiers

Disulfide-bond 1..11 /note="This bond cyclises the peptide"

WO200152875-A1.

26-JUL-2001.

18-JAN-2001; 2001WO-US01533.

18-JAN-2000; 2000US-0176293.

16-MAY-2000; 2000US-0204590.

(LUDW-) LUDWIG INST CANCER RES.

Achen MG, Hughes RA, Stacker S, Cendron A;

DR WPI, 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

PS Claim 49, Page 32, 102pp: English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC a beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthritis.
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heart or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

SQ Sequence 11 AA:

Query Match

Best Local Similarity 100.0%; Score 61; DB 22; Length 11;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTSVPLTSVPC 11

DB 1 CTSVPLTSVPC 11

RESULT 2

ID AA004542 standard; Peptide; 11 AA.

AC AA004542;

DT 26-SEP-2001 (first entry)

DE VEGF based monocyclic peptide 20.

KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

Key Location/Qualifiers

FT Disulfide-bond 1..11 /note="This bond cyclises the peptide"

PN WO200152875-A1.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US01533.

XX 18-JAN-2000; 2000US-0176293.

XX 16-MAY-2000; 2000US-0204590.

XX (LUDM-) LUDMIG INST CANCER RES.

XX Achen MG, Hughes RA, Stacker S, Cendron A;

XX WPI, 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

PS Example 25, Page 47, 102pp: English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC a beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthritis.
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heart or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

SQ Sequence 11 AA:

Query Match

Best Local Similarity 96.7%; Score 59; DB 22; Length 11;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTSVPLTSVPC 11

DB 1 CTSVPLTSVPC 11

RESULT 3

ID AA004545 standard; Peptide; 11 AA.

AC AA004545;

DT 26-SEP-2001 (first entry)

DE VEGF based monocyclic peptide 23.

KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.

XX OS Synthetic.
XX XX Location/Qualifiers
FT FT Key 1..11
FT FT Disulfide-bond /note="This bond cyclises the peptide"
XX XX
XX PN WO200152875-A1.
XX PD 26-JUL-2001.
XX PF 18-JAN-2001; 2001WO-US01533.
XX PR 18-JAN-2000; 2000US-0176293.
XX PR 16-MAY-2000; 2000US-0204590.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX PI WPI; 2001-442248/47.
XX DR
XX XX
XX PT Novel monomeric monocyclic peptide, used to interfere with
XX PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
XX PT loop fragment from an exposed loop of a growth factor protein by
XX PT oxidizing the cysteine residues -
XX PS
XX PS Example 25; Page 47; 102pp; English.
XX XX
XX CC The sequence represents a monomeric monocyclic peptide of the invention,
XX CC whose 3-dimensional structure is modelled on the expose loop of human
XX CC VEGF (vascular endothelial growth factor). The invention relates to a
XX CC method of producing a monomeric monocyclic peptide by a measuring
XX CC beta-beta carbon separation distances on opposite antiparallel strands of
XX CC a peptide loop fragment from an exposed loop of a growth factor protein
XX CC and cyclising the peptide by oxidising the cysteine residues. The
XX CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX CC monocyclic peptides) and a cyclic peptide with at least one amino acid
XX CC deleted prior to cyclisation are used to interfere with angiogenesis,
XX CC neovascularisation or lymphangiogenesis in a mammal with a condition
XX CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX CC The condition is diabetic retinopathy, psoriasis, arthropathy,
XX CC hemangioma, vascularised malignant or benign tumour, post-recovery
XX CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX CC trauma, substance-induced neovascularisation of the liver, excessive
XX CC hormone-related angiogenic dysfunction, diabetes induced neovascular
XX CC sequelae, hypertension induced neovascular sequelae, or chronic liver
XX CC infection. The peptides are also used to modulate vascular permeability
XX CC in a mammal (the mammal has a condition characterised by fluid
XX CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX CC or brain. The peptides are used to image blood vessels and lymphatic
XX CC vasculature. The monomeric and bicyclic peptides are used to interfere
XX CC with at least one biological activity induced by VEGF, VEGF-C or -D and
XX CC are also used in combination with an anti-inflammatory agent, to treat a
XX CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
XX CC diabetic retinopathy.
XX SQ Sequence 11 AA;
QY Query Match 93.4%; Score 57; DB 22; Length 11;
DB Best Local Similarity 81.8%; Pred. No. 0.0034; 0; Indels 0; Gaps 0;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTSVPLTSPVC 11
DB 1 CTSVPLTSPVC 11

RESULT 4
AAU04543
ID AAU04543 standard; Peptide; 11 AA.
XX

AC AAU04543;
XX XX 26-SEP-2001 (first entry)
XX XX VEGF based monocyclic peptide 21.
XX DE
XX XX Human; VEGF; vascular endothelial growth factor; angiogenesis;
XX KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
XX KW diabetes induced neovascular sequelae; rheumatoid arthritis;
XX KW diabetic retinopathy; chronic inflammation; cyclic.
XX OS
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FH Disulfide-bond 1..11
XX FT /note="This bond cyclises the peptide"
XX FT
XX PN WO200152875-A1.
XX PD 26-JUL-2001.
XX PF 18-JAN-2001; 2001WO-US01533.
XX PR 18-JAN-2000; 2000US-0176293.
XX PR 16-MAY-2000; 2000US-0204590.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX PI WPI; 2001-442248/47.
XX DR
XX XX
XX PT Novel monomeric monocyclic peptide, used to interfere with
XX PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
XX PT loop fragment from an exposed loop of a growth factor protein by
XX PT oxidizing the cysteine residues -
XX PS
XX PS Example 25; Page 47; 102pp; English.
XX XX
XX CC The sequence represents a monomeric monocyclic peptide of the invention,
XX CC whose 3-dimensional structure is modelled on the expose loop of human
XX CC VEGF (vascular endothelial growth factor). The invention relates to a
XX CC method of producing a monomeric monocyclic peptide by a measuring
XX CC beta-beta carbon separation distances on opposite antiparallel strands of
XX CC a peptide loop fragment from an exposed loop of a growth factor protein
XX CC and cyclising the peptide by oxidising the cysteine residues. The
XX CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX CC monocyclic peptides) and a cyclic peptide with at least one amino acid
XX CC deleted prior to cyclisation are used to interfere with angiogenesis,
XX CC neovascularisation or lymphangiogenesis in a mammal with a condition
XX CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX CC The condition is diabetic retinopathy, psoriasis, arthropathy,
XX CC hemangioma, vascularised malignant or benign tumour, post-recovery
XX CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX CC trauma, substance-induced neovascularisation of the liver, excessive
XX CC hormone-related angiogenic dysfunction, diabetes induced neovascular
XX CC sequelae, hypertension induced neovascular sequelae, or chronic liver
XX CC infection. The peptides are also used to modulate vascular permeability
XX CC in a mammal (the mammal has a condition characterised by fluid
XX CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX CC or brain. The peptides are used to image blood vessels and lymphatic
XX CC vasculature. The monomeric and bicyclic peptides are used to interfere
XX CC with at least one biological activity induced by VEGF, VEGF-C or -D and
XX CC are also used in combination with an anti-inflammatory agent, to treat a
XX CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
XX CC diabetic retinopathy.
XX SQ Sequence 11 AA;
QY Query Match 88.5%; Score 54; DB 22; Length 11;
DB Best Local Similarity 72.7%; Pred. No. 0.011;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTSVPLTSPVC 11
DB 1 CTSVPLTSPVC 11

OY 1 C1SVPLTSPVC 11
 11:111111
 DB 1 c1svpltsipc 11

RESULT 5

AAU04544
 ID AAU04544 standard; Peptide: 11 AA.

AC AAU04544;

DT 26-SEP-2001 (first entry)

DE VEGF based monocyclic peptide 22.

Human: VEGF; vascular endothelial growth factor; angiogenesis;
 neovascularisation; lymphangiogenesis; psoriasis; tumour;
 diabetes induced neovascular sequelae; rheumatoid arthritis;
 diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

Key Location/Qualifiers

FT Disulfide-bond 1..11

FT /note="This bond cyclises the peptide"

PN WO200152875-A1.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01533.

PR 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

RA (LUDW-) LUDWIG INST CANCER RES.

PI Achen MG, Hughes RA, Stacker S, Cendron A;

DR WPI; 2001-442248/47.

PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -

PS Example 25; Page 47; 102pp; English.

CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthritis,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and

CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.

XX Sequence 11 AA;

Query Match 85.2%; Score 52; DB 22; Length 11;
 Best Local Similarity 72.7%; Pred. NO. 0.023; Indels 0; Gaps 0;
 Matches 8; Conservative 3; Mismatches 0;

OY 1 C1SVPLTSPVC 11
 11:111111
 DB 1 c1svpltsipc 11

RESULT 6

AAU04532
 ID AAU04532 standard; Peptide: 10 AA.

AC AAU04532;

DT 26-SEP-2001 (first entry)

DE VEGF based monocyclic peptide 10.

Human: VEGF; vascular endothelial growth factor; angiogenesis;
 neovascularisation; lymphangiogenesis; psoriasis; tumour;
 diabetes induced neovascular sequelae; rheumatoid arthritis;
 diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

Key Location/Qualifiers

FT Disulfide-bond 1..10

FT /note="This bond cyclises the peptide"

PN WO200152875-A1.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01533.

PR 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

RA (LUDW-) LUDWIG INST CANCER RES.

PI Achen MG, Hughes RA, Stacker S, Cendron A;

DR WPI; 2001-442248/47.

PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -

PS Claim 49; Page 32; 102pp; English.

CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthritis,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery

CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D are used to
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D can
CC be useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
CX

50 Sequence 109 AA;

```

Query Match      70.5%; Score 43; DB 20; Length 109;
Best Local Similarity 100.0%; Prid. No. 7-7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ISVPLTSVP 10
    |||||
Db 76 ISVPLTSVP 84

```

RESULT	9
AAB11931	

1D AAB11931 standard; protein; 109 AA.
XX

AC AAB11931;
YY

20-NOV-2000 (first entry)

Human truncated VEGF-D.

11uncate; Egr-B; vascular endothelial growth factor; human;
12monoclonal antibody; VEGF receptor; VEGFR-2; VEGFR-3;
13vascular permeability disorder; endothelial cell proliferative disorder
14angiogenic disorder; lymphangiogenic disorder;
15neovascularization disorder; endothelial cell differentiation disorder;
16cancer; diabetic retinopathy; psoriasis; arthropathy; pulmonary oedema;
17detection; diagnosis; imaging; lymphatic vasculature.

Homo sapiens.

WO2000037025-A2

29-JUN-2000

21-DEC-1999; 99WO-US31332.

21-DEC-1998; 98US-0113254

4/-MAY-1999; 9905-0134556.

(LUDW-) LUDWIG INST CANCER RES.

Achen MG, Stacker SA;

WPI; 2000-442498/38

novel compositions comprising antipodics reactive to vascular endothelial growth factor-D, useful for treating, e.g. angiogenesis, lymphangiogenesis and neovascularization disorders -

claim 1; Fig 1; 44pp; English.

C This sequence represents a 100

CC these receptors and additionally is not reactive with VEGF-C. The
CC fragments thereof which is specifically reactive with the truncated
CC human VEGF-1 and method of preparing the antibody. The antibody of the
CC invention interferes with the binding of VEGF-D to the VEGF receptors
CC VEGFR-2 and VEGFR-3 but does not interfere with the binding of VEGF to
CC these receptors and additionally is not reactive with VEGF-C. The

CC antibody may be used to treat disorders associated with vascular
CC permeability, endothelial cell proliferation, angiogenesis,
CC lymphangiogenesis, neovascularisation and endothelial cell
CC differentiation, especially cancer, diabetic retinopathy, psoriasis, and
CC arthropathies. The antibody may also be used to treat fluid accumulation
CC in the heart and/or lung via modulation of vascular permeability. It may
CC additionally be used to detect VEGF-D and may be used to image lymphatic
CC vasculature in tissue.

SQ **Sequence** **109 AA;**

Query Match	70.5%;	Score 43;	DB 21;	Length 109;
Best Local	Similarity 100.0%;	Pred. No. 7.7;		
Matches	9;	Conservative 0;	Mismatches 0;	Indels 0;
QY	2	ISVPLTSVP	10	
Db	76	isvpltsvp	84	

RESULT 10
AAY08287

ID	standard; Protein; 178 AA.
XX	

AC AY08287;
XY

14-JUL-1999 (first entry)

Human growth factor protein fragment FIGF178 (VEGF-D178).

pharmaceutical.

Homo sapiens.

DE19748734-A1.

06-MAY-1999

05-NOV-1997; 97DE-1048734.

05-NOV-1997; 97DE-1048734.

(GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH

Erdmann H, Kaerst U, Mueller C, Rinas U, Welch H

WPI; 1999-278785/24

Preparing active growth factor dimers from inclusion bodies in high yield

This invention describes the no

The invention describes the novel preparation of biologically active dimers of recombinant human growth factors of the cysteine knot family starting from cellular inclusion bodies. Such dimers are useful in pharmaceutical compositions and the method provides yields of 31-39.7% in examples, compared with about 10% for the conventional method. (see Biochemistry, 28 (1989) 2956). AAY08278-Y08301 are human growth factor protein fragments used in the method of the invention.

Sequence 178 AA;

	Query Match	Score 43;	DB 20;	Length 178;
	Best Local Similarity	100.0%;	Pred No. 13;	
Matches	9;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;
2	ISVPLTSP	10		
153	isvpltsvp	161		

RESULT 11
ID AAM53243 standard; Protein; 321 AA.
XX
AC AAM53243;
XX
DF 03-AUG-1998 (first entry)
XX
DE Mus musculus vascular endothelial growth factor D2 (VEGF-D2).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Mus musculus.
XX
PN W09807832-A1.
XX
PD 26-FEB-1998.
XX
PE 21-AUG-1997; 97WO-US14696.
XX
PR 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Achen MG, Alltalo K, Stacker SA, Wilks AF;
XX
DR WPI: 1998-179057/16.
DR N-PSDB: AAV20809.
XX
PT New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
PS Claim 16: Pages 64-65; 101pp; English.
XX
The sequence is that of mouse lung vascular endothelial growth factor
D2 (VEGF-D2). VEGF-D2 can be used for e.g. acceleration of angiogenesis
in wound healing, tissue or organ transplantation, or to establish
collateral circulation in tissue infarction or arterial stenosis,
such as coronary artery disease, and inhibition of angiogenesis in
the treatment of cancer or of diabetic retinopathy. It can also be
used in the treatment of lung disorders to improve blood circulation
in the lung and/or gaseous exchange between the lungs and the blood
stream or to improve blood circulation to the heart and O2 gas
permeability in cases of cardiac insufficiency, to improve blood
flow and gaseous exchange in chronic obstructive airway disease,
or to treat malabsorptive syndromes in the intestinal tract.
Quantitation of VEGF-D in cancer biopsy specimens may be useful
as an indicator of future metastatic risk. Antagonists can be used
for treating e.g. conditions such as congestive heart failure,
involving accumulations of fluid in the lung resulting from
increases in vascular permeability. The products can also be used
for detection and diagnosis.

Query Match 70.5%; Score 43; DB 19; Length 321;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 ISVPLTSVP 10
DB 168 ISVPLTSVP 176
RESULT 12
ID AAM53240 standard; Protein; 325 AA.
XX
AC AAM53240;
XX
DF 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH 126..128
FH /note="potential N-linked glycosylation site"
FH 156..158
FH /note="potential N-linked glycosylation site"
FH 258..260
FH /note="potential N-linked glycosylation site"
XX
PN W09807832-A1.
XX
PD 26-FEB-1998.
XX
PE 21-AUG-1997; 97WO-US14696.
XX
PR 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Achen MG, Alltalo K, Stacker SA, Wilks AF;
XX
DR WPI: 1998-179057/16.
DR N-PSDB: AAV20806.
XX
PT New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
PS Claim 16: Pages 57-58; 101pp; English.
XX
The sequence is that of human breast vascular endothelial growth factor
D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
in wound healing, tissue or organ transplantation, or to establish
collateral circulation in tissue infarction or arterial stenosis,
such as coronary artery disease, and inhibition of angiogenesis in

CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O₂ gas
CC permeability in cases of cardiac insufficiency, to improve blood
CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.

SO Sequence 325 AA;

Query Match 70.5%; Score 43; DB 19; Length 325;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
| | | | | | | | | |
Db 139 isvpltsvp 147

RESULT 13
AA97572
ID AAY97572 standard; Protein; 325 AA.

AC AAY97572;

DT 05-APR-2001 (first entry)

DE Human VEGF-D protein sequence.

XX Human; angiogenic protein; wound healing; vascular tissue repair;
KW peripheral arterial disease; critical limb ischemia; coronary disease;
KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KW infectious disease; neurodegeneration;
KW vascular endothelial growth factor-D; VEGF-D.

OS Homo sapiens.

PN WO200075163-A1.

PD 14-DEC-2000.

PF 01-JUN-2000; 2000WO-US14925.

PR 03-JUN-1999; 99US-0137796.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Hu J, Cao L;

DR WPI; 2001-071057/08.

DR N-PSDB; AAA91006.

PT New nucleic acid encoding angiogenic proteins, useful e.g. for
PT promoting healing of wounds and treating peripheral arterial disease,
PT critical limb ischemia or coronary disease -

PS Claim 11; Page 226-227; 244pp; English.

CC This sequence is vascular endothelial growth factor-D (VEGF-D),
CC which is an angiogenic protein of the invention. The angiogenic proteins
CC and the DNA sequences encoding them, are used to prevent, treat or
CC ameliorate disease and to detect diseases, or susceptibility, by
CC detecting mutations or the presence or amount of angiogenic protein
CC expression. Particularly they are used to stimulate wound healing,
CC growth of damaged bone and tissue, and for repair of vascular tissue,

CC especially peripheral arterial disease, critical limb ischemia or
CC coronary disease. Antagonists of the sequences are used to inhibit
CC angiogenesis in tumours and to treat inflammation (where associated with
CC increased vascular permeability), diabetic retinopathy, rheumatoid
CC arthritis or psoriasis. Agonists are also useful for stimulating
CC (lymph)angiogenesis. The proteins are also used to identify specific
CC binding agents (potential therapeutic agents) and to raise antibodies.
CC The antibodies are useful as therapeutic (ant)agonists; for detection,
CC purification and targeting of proteins for in vivo or in vitro diagnosis
CC (including imaging) or for therapy (including when linked to e.g. a label
CC or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
CC residual disease or haematopoietic progenitor/stem cells. It is also
CC contemplated that the sequences might be useful for treating a very wide
CC range of other disorders, e.g. autoimmune diseases; allergy; cancer;
CC infectious diseases (viral, bacterial, fungal or parasitic);
CC neurodegeneration, also as chemotactic agents or for stimulating
CC regeneration of the nervous system etc.

SO Sequence 325 AA;

Query Match 70.5%; Score 43; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
| | | | | | | | | |
Db 139 isvpltsvp 147

RESULT 14
AA44296
ID AAM44296 standard; Protein; 326 AA.

AC AAM44296;

DT 22-JUN-1998 (first entry)

DE Rat vascular endothelial growth factor D.

XX Rate; vascular endothelial growth factor D; VEGF-D; gene therapy;
KW inflammation; oedema.

OS Rattus sp.

PN WO9802543-A1.

PD 22-JAN-1998.

PF 15-JUL-1997; 97WO-JP02456.

PR 15-JUL-1996; 96JP-0185216.

PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PI Hirata Y, Nezu J;

DR WPI; 1998-110591/10.

DR N-PSDB; AAV15178.

PT VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and
PT treating oedema

PS Example 8; Page 35-38; 52pp; Japanese.

CC The present sequence represents rat vascular endothelial growth factor D
CC (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind
CC the protein, may be useful in, e.g. gene therapy and in treatment of
CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D
CC DNA sequences may be used for screening for the compounds which bind to
CC the VEGF-D protein.

SO Sequence 326 AA;

Query Match 70.5%; Score 43; DB 19; Length 326;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
 |||||
 DB 173 isvpltsvp 181

RESULT 15

AA08286
 ID AA08286 standard; Protein; 337 AA.

AC AA08286;

DT 14-JUL-1999 (first entry)

DE Human growth factor protein fragment FIGF (VEGF-D).

XX Growth factor; human; dimer; cysteine knot; cellular inclusion body;
 KW pharmaceutical.

XX Homo sapiens.

OS DE19748734-A1.

PN 06-MAY-1999.

PD 05-NOV-1997; 97DE-1048734.

PR 05-NOV-1997; 97DE-1048734.

PA (GBFB) GBS BIOTECHNOLOGISCHE FORSCHUNG MBH.

PI Erdmann H, Kaerst U, Mueller C, Rinas U, Welch H;

DR WPI: 1999-278785/24.

XX Preparing active growth factor dimers from inclusion bodies in high
 PT yield

PS Claim 14; Page 9; 14pp; German.

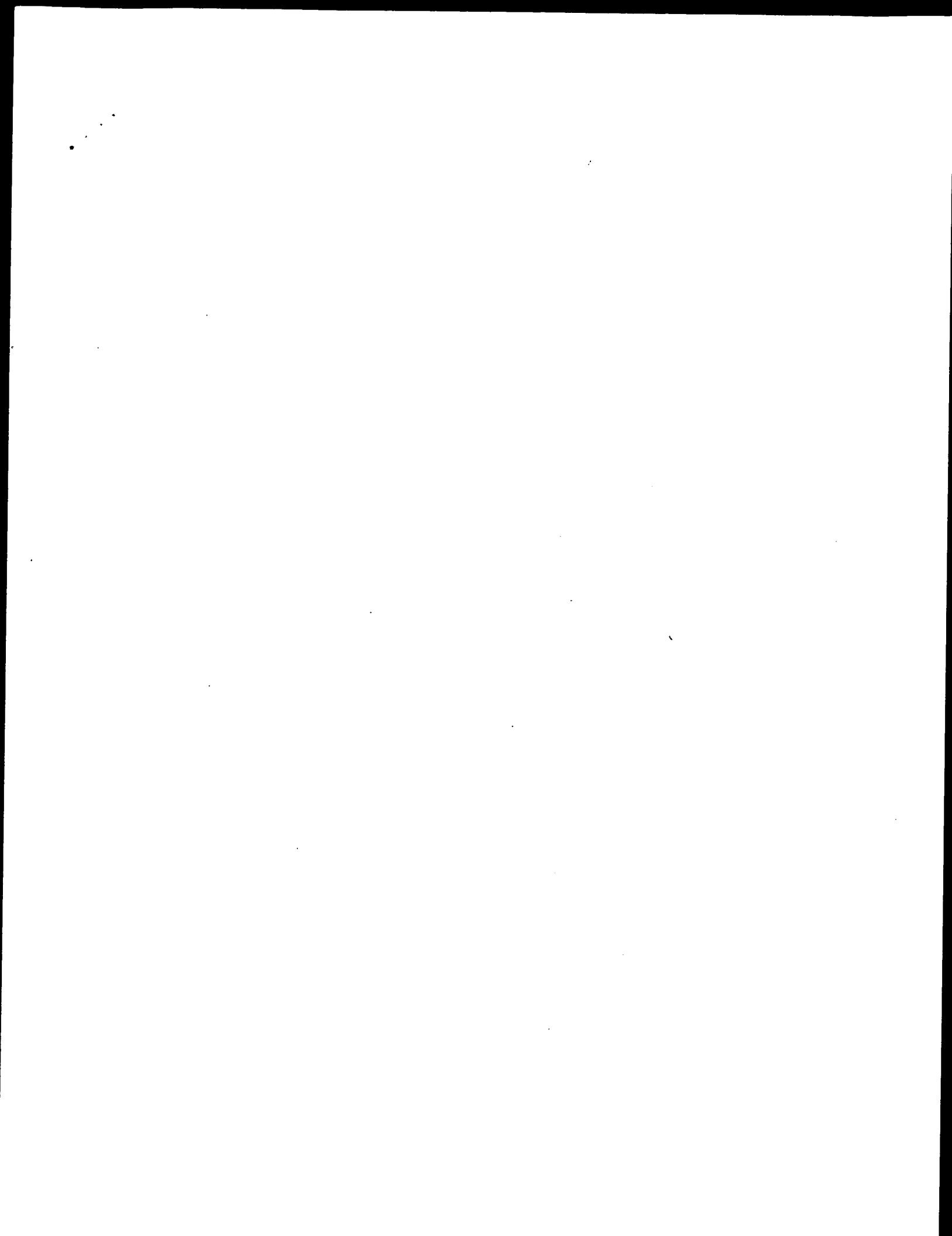
XX This invention describes the novel preparation of biologically active
 CC dimers of recombinant human growth factors of the cysteine knot family
 CC starting from cellular inclusion bodies. Such dimers are useful in
 CC pharmaceutical compositions and the method provides yields of 31-39.7%,
 CC in examples, compared with about 10% for the conventional method (see
 CC Biochemistry, 28 (1989) 2956). AA08278-Y08301 are human growth factor
 CC protein fragments used in the method of the invention.

SQ Sequence 337 AA;

Query Match 70.5%; Score 43; DB 20; Length 337;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
 |||||
 DB 152 isvpltsvp 160

Search completed: June 17, 2002, 16:02:12
 Job time: 421 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:36 ; Search time 75.98 Seconds
(without alignments)
3.536 Million cell updates/sec

Title: US-09-761-636A-7
Perfect score: 61
Sequence: 1 CISVPLTSPVC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents, AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	43	70.5	321 4	US-08-915-795-9 Sequence 9, Appli
2	43	70.5	325 4	US-08-915-795-3 Sequence 3, Appli
3	43	70.5	354 4	US-08-915-795-5 Sequence 5, Appli
4	43	70.5	358 4	US-08-915-795-8 Sequence 8, Appli
5	38	62.3	663 4	US-08-959-004-5 Sequence 5, Appli
6	36	59.0	388 4	US-08-861-774E-94 Sequence 94, Appli
7	35	57.4	90 6	5220013-24 Patent No. 5220013
8	35	57.4	208 4	US-08-612-973-32 Sequence 32, Appli
9	35	57.4	208 4	US-08-927-597-32 Sequence 32, Appli
10	35	57.4	299 2	US-08-923-856-1 Sequence 1, Appli
11	35	57.4	299 3	US-09-216-294-1 Sequence 1, Appli
12	35	57.4	363 1	US-08-484-105-20 Sequence 20, Appli
13	35	57.4	363 1	US-08-484-106-20 Sequence 20, Appli
14	35	57.4	22 4	US-09-227-557-517 Sequence 517, App
15	34	55.7	38 1	US-08-176-500-111 Sequence 111, App
16	34	55.7	38 1	US-08-471-052A-111 Sequence 111, App
17	34	55.7	38 1	US-08-189-331-111 Sequence 111, App
18	34	55.7	38 2	US-08-471-939-111 Sequence 111, App
19	34	55.7	38 2	US-08-471-800-111 Sequence 111, App
20	34	55.7	38 2	US-08-471-068-111 Sequence 111, App
21	34	55.7	72 3	US-08-722-719-58 Sequence 58, Appli
22	34	55.7	73 3	US-08-722-719-26 Sequence 26, Appli
23	34	55.7	73 3	US-08-722-719-57 Sequence 57, Appli
24	34	55.7	75 3	US-08-722-719-55 Sequence 55, Appli
25	34	55.7	76 3	US-08-722-719-24 Sequence 24, Appli
26	34	55.7	76 3	US-08-722-719-56 Sequence 56, Appli
27	34	55.7	76 3	US-08-722-719-59 Sequence 59, Appli

28	34	55.7	77 3	US-08-722-719-22 Sequence 22, Appli
29	34	55.7	77 3	US-08-722-719-61 Sequence 61, Appli
30	34	55.7	77 4	US-09-044-796A-12 Sequence 12, Appli
31	34	55.7	78 3	US-08-722-719-62 Sequence 62, Appli
32	34	55.7	82 3	US-08-722-719-60 Sequence 60, Appli
33	34	55.7	83 3	US-08-722-719-18 Sequence 18, Appli
34	34	55.7	84 3	US-08-722-719-20 Sequence 20, Appli
35	34	55.7	100 3	US-08-722-719-30 Sequence 30, Appli
36	34	55.7	100 3	US-08-722-719-32 Sequence 32, Appli
37	34	55.7	100 3	US-08-722-719-54 Sequence 54, Appli
38	34	55.7	120 3	US-08-722-719-4 Sequence 4, Appli
39	34	55.7	120 4	US-08-681-192-6 Sequence 6, Appli
40	34	55.7	121 1	US-08-208-339A-2 Sequence 2, Appli
41	34	55.7	137 2	US-08-421-144A-2 Sequence 2, Appli
42	34	55.7	1481 2	US-08-616-844-40 Sequence 40, Appli
43	34	55.7	1481 2	US-08-599-654-40 Sequence 40, Appli
44	34	55.7	1481 3	US-08-944-668A-40 Sequence 40, Appli
45	34	55.7	1481 3	US-08-944-423A-40 Sequence 40, Appli

ALIGNMENTS

RESULT 1
US-08-915-795-9
Sequence 9, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Karl ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET INFORMATION: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-9
Query Match 70.5%; Score 43; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Gaps 0;

OY 2 ISVPLTSVP 10
|||||||
Db 168 ISVPLTSVP 176

RESULT 2

US-08-915-795-3
; Sequence 3, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Breast
; US-08-915-795-3

Query Match

Best Local Similarity 70.5%; Score 43; DB 4; Length 325;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
|||||||
Db 139 ISVPLTSVP 147

RESULT 3

US-08-915-795-5
; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 354 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

TISSUE TYPE: Human Lung

US-08-915-795-5

Query Match

Best Local Similarity 70.5%; Score 43; DB 4; Length 354;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
|||||||
Db 168 ISVPLTSVP 176

RESULT 4

US-08-915-795-8
; Sequence 8, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-8

Query Match 70.5%; Score 43; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSP 10
|||||
Db 173 ISVPLTSP 181

RESULT 5
US-08-959-004-5
Sequence 5, Application US/08959004
Patent No. 6197543
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Puri
APPLICANT: Kaser, Matthew
TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,004
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0414 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 663 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ADREUT06
CLONE: 2822412
US-08-959-004-5

Query Match 62.3%; Score 38; DB 4; Length 663;
Best Local Similarity 77.8%; Pred. No. 1,4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CISVPLTSV 9
|||||
Db 479 CISVPLTET 487

RESULT 6
US-08-861-774E-94
Sequence 94, Application US/08861774E
Patent No. 6297007
GENERAL INFORMATION:
APPLICANT: Waters, Barbara
APPLICANT: Miao, Vivian
APPLICANT: Ho, Yap
APPLICANT: Tong, Seow
TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR
FILE REFERENCE: 9993-006
CURRENT APPLICATION NUMBER: US/08/861,774E
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 94
LENGTH: 388
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Clone ps3
US-08-861-774E-94

Query Match 59.0%; Score 36; DB 4; Length 388;
Best Local Similarity 77.8%; Pred. No. 1,6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ISVPLTSP 10
|||||
Db 318 YVPLTSP 326

RESULT 7
US-08-952-0013-24
Patent No. 5220013
APPLICANT: PONTE, PHYLLIS A.;CORDELL, BARBARA
TITLE OF INVENTION: DNA SEQUENCE USEFUL FOR THE DETECTION
OF ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/444,118
FILING DATE: 30-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 87,002
FILING DATE: 18-AUG-1987
APPLICATION NUMBER: 8,810
FILING DATE: 30-JAN-1987
APPLICATION NUMBER: 948,376
FILING DATE: 31-DEC-1986
APPLICATION NUMBER: 932,193
FILING DATE: 17-NOV-1986
SEQ ID NO:24:

LENGTH: 90
5220013-24

Query Match 57.4%; Score 35; DB 6; Length 90;
Best Local Similarity 45.5%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 C1SVPLTSPC 11
DB 19 CLEPPYGCPC 29

RESULT 8
US-08-612-973-32

Sequence 32, Application US/08612973
Patent No. 6150134

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS

APPLICANT: DE MARTYNOFF, GUY

TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHVE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/612,973

FILING DATE: 11-MAR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4100

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 208 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-612-973-32

QY 1 C1SVPLTSP 10
DB 66 CLTVPASAVP 75

RESULT 9

US-08-927-597-32

Sequence 32, Application US/08927597

Patent No. 6245503

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS

APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BOYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHVE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/927,597

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/612,973

FILING DATE: 11-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4100

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 208 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-927-597-32

QY 1 C1SVPLTSP 10
DB 66 CLTVPASAVP 75

RESULT 10
US-08-923-856-1

Sequence 1, Application US/08923856

Patent No. 5928894

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Tang, Tom

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN ACTVA-ORF4-LIKE PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/923,856
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0380 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCNOT01
CLONE: 223909
US-08-923-856-1

Query Match 57.4%; Score 35; DB 2; Length 299;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 2 ISVPLTSV--PC 11
1111111111
Db 143 IGVPMTSVRLPC 154

RESULT 11
US-09-216-294-1
Sequence 1, Application US/09216294
Patent No. 6080723
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Tang, Tom
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Puri
TITLE OF INVENTION: HUMAN ACTIVA-ORF4-LIKE PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/216,294
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/923,856
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0380 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCNOT01
CLONE: 223909
US-09-216-294-1

Query Match 57.4%; Score 35; DB 3; Length 299;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 2 ISVPLTSV--PC 11
1111111111
Db 143 IGVPMTSVRLPC 154

RESULT 12
US-08-484-105-20
Sequence 20, Application US/08484105
Patent No. 5589341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELT, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 363 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-105-20

Query Match 57.4%; Score 35; DB 1; Length 363;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CISVPLTS 8
1111111111

DB 344 CLNIPLTS 351

RESULT 13

US-08-484-106-20

Sequence 20, Application US/08484106

Patent No. 5614618

GENERAL INFORMATION:

APPLICANT: STILLMAN, Bruce

APPLICANT: BELL, Stephen P

APPLICANT: KOBAYASHI, Ryuji

APPLICANT: RINE, Jasper

APPLICANT: FOSS, Margit

APPLICANT: MCNALLY, Francis J

APPLICANT: LAURENSEN, Patricia

APPLICANT: HERSKOWITZ, Ira

APPLICANT: LI, Joachim J

APPLICANT: GAVIN, Kimberly

TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,106

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman Ph.D., Richard Aron

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 363 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-484-106-20

Query Match 57.4%; Score 35; DB 1; Length 363;
Best local similarity 62.5%; Pred No. 2.2e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVP LITS 8

Db 344 CLNIP LITS 351

RESULT 14

US-09-227-357-517

Sequence 517, Application US/09227357

Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1

CURRENT APPLICATION NUMBER: US/09/227,357

EARLIER FILING DATE: 1999-01-08

EARLIER APPLICATION NUMBER: PCT/US98/13684

EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 517
LENGTH: 22
TYPE: PRT
ORGANISM: Homo sapiens
US-09-227-357-517

Query Match 55.7%; Score 34; DB 4; Length 22;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLTSVP 10
1 1 1 1 1 1
DB 6 C1SVPLGSCP 15

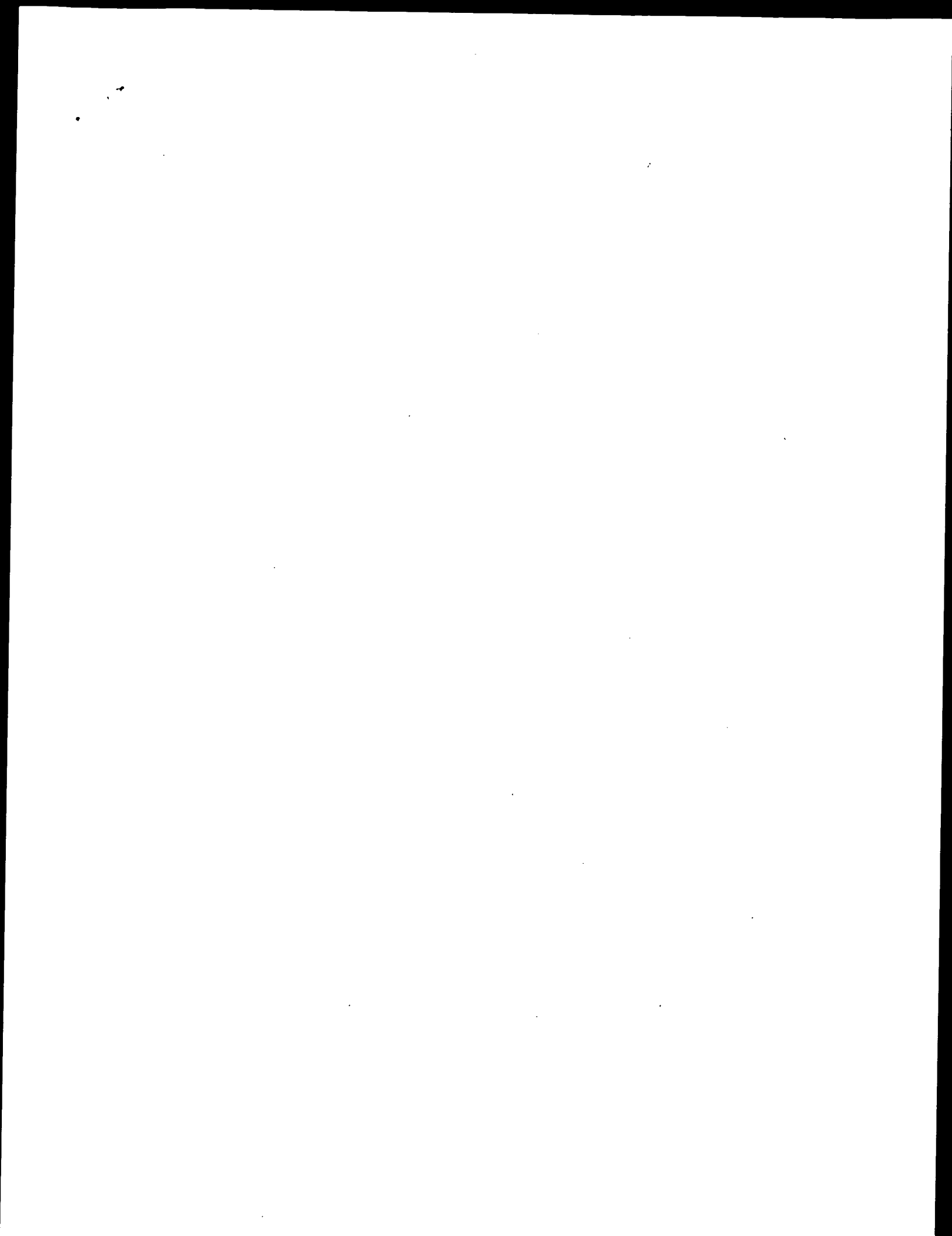
RESULT 15

US-08-176-500-111
; Sequence 111, Application US/08176500
; Patent No. 5498538
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-176-500-111

Query Match 55.7%; Score 34; DB 1; Length 38;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLTSVP 10
1 1 1 1 1 1
DB 3 LTLPLTSIP 11

Search completed: June 17, 2002, 15:56:36
Job time: 85 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:21 ; Search time 95.43 Seconds
(without alignments)
17.117 Million cell updates/sec

Title: US-09-761-636a-8
Perfect score: 100
Sequence: 1 CASLGKSTNFCCKPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58.0	419	2	S69207	vascular endothelial
2	48.0	211	1	XUBS0M	methyphosphotrieth
3	48.0	330	1	T23169	hypothetical prote
4	45.0	371	2	B96614	hypothetical prote
5	44.0	1717	1	A45558	epidermal growth f
6	43.0	100	2	D71882	hypothetical prote
7	43.0	104	2	A39035	ribonuclease-relat
8	43.0	693	2	S64904	probable membrane
9	42.0	1428	2	T08852	lustrin A - Califo
10	41.0	223	2	B38346	ultra-high-sulfur
11	41.0	330	2	A38346	ultra-high-sulfur
12	41.0	319	2	A84947	lytB protein (Impo
13	41.0	335	2	B45511	chitinase (EC 3.2.
14	41.0	397	2	AC1570	acetate kinase hom
15	41.0	397	2	AH1220	acetate kinase hom
16	41.0	406	2	T30748	hypothetical prote
17	41.0	439	2	S33293	testican - human
18	41.0	476	2	F88968	protein T27B7.2 (l
19	41.0	570	2	T37314	probable Kexin (EC
20	41.0	585	2	C69336	probable electron
21	41.0	642	1	QOBER5	UL77 protein - hum
22	41.0	669	2	S65551	factor H - bovine
23	41.0	815	2	T05555	DNA polymerase III
24	41.0	942	2	D87803	protein b13-4D (Im
25	41.0	988	1	S35362	protein kinase C (
26	41.0	1174	2	T43051	probable membrane
27	41.0	1294	2	S77690	hypothetical prote
28	40.0	1294	2	E97808	hypothetical prote
29	40.0	126	2	S30858	hypothetical prote

ALIGNMENTS

RESULT 1

S69207 vascular endothelial growth factor C precursor - human

N:Alternate names: FLT4 ligand DHM

C:Species: Homo sapiens (man)

C>Date: 27-Apr-1996 #sequence-revision 01-Nov-1996 #text-change 08-Oct-1999

C:Accession: S69207; S61795; S71443; S69208; G02659

R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Sa

EMBO J. 15, 1751, 1996

A>Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a li

A:Reference number: S69207; M01D:96203094

A:Accession: S69207

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-419 <JOU>

A:Cross-references: EMBL:X94216; NID:q1177488; PIDN:CA63907.1; PID:e221096; PID:

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December

A>Note: only a part of the translation is shown

A>Note: this is a revision to the sequence from reference S61795

R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Sa

EMBO J. 15, 290-298, 1996

A>Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the

A:Reference number: S61795; M01D:96178224

A:Accession: S61795

A>Status: nucleic acid sequence not shown; not compared with conceptual translati

A:Molecule type: mRNA

A:Residues: 70-419 <JOU1>

A>Note: this sequence has been revised in reference S69207

A:Accession: S71443

A:Molecule type: protein

A:Residues: 'X',104-120 <JOU2>

R:Lee, J.; Gray, A.; Yuan, J.; Luoh, S.M.; Avraham, H.; Wood, W.I.

submitted to the EMBL Data Library, December 1995

A:Description: Vascular endothelial growth factor related protein (VRP): A ligand

A:Reference number: S69208

A:Accession: S69208

A:Molecule type: mRNA

A:Residues: 1-419 <LEB>

A:Cross-references: EMBL:U43142; NID:q1150988; PIDN:AAA65214.1; PID:q1150989

R:Morris, J.C.

submitted to the EMBL Data Library, May 1996

A:Reference number: H01557

A:Accession: G02659

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-419 <MOR>

A:Cross-references: EMBL:U58111; NID:q1373426; PIDN:AAB02909.1; PID:q1373427

F:103-419/Product: vascular endothelial growth factor C #status experimental <MAT>

Query Match 58.0%; Score 58; DB 2; Length 419;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ELKSTNTFCCKPCC 17
Db 143 EFGVATNTFFCKPCC 156

RESULT 2

methylnphosphotriester-DNA alkyltransferase (EC 2.1.1.-) / adaAB operon transcription act
C:Species: Bacillus subtilis
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jun-2000
C:Accession: S11483; F69582
R:Morohoshi, F.; Hayashi, K.; Munakata, N.
Nucleic Acids Res. 18, 5473-5480, 1990
A:Title: Bacillus subtilis ada operon encodes two DNA alkyltransferases.
A:Reference number: S11483; MUID:91016831
A:Accession: S11483
A:Molecule type: DNA
A:Residues: 1-211 <MOR>

A:Cross-references: EMBL:X53399; NID:939786; PIDN:CA37475.1; PID:939787
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galliz, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Luita, K.; Lapidus, A.; Lardinois,
Y., M.; Ogawa, K.; Ogilvie, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rochoa, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroefer, F.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
T.; Whiners, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033
A:Accession: F69582
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-211 <KCN>
A:Cross-references: GB:299104; GB:299105; GB:AL009126; NID:92632457; PIDN:CAB11974.1; PI
A:Experimental source: strain 168
C:Genetics:
A:Gene: adaA

C:Superfamily: methylnphosphotriester-DNA methyltransferase; methylnphosphotriester-DNA me
C:Keywords: DNA binding; DNA repair; methyltransferase; transcription regulation
F:23-207/Domain: methylnphosphotriester-DNA methyltransferase homology <MPT>

Query Match 48.0%; Score 48; DB 1; Length 211;
Best Local Similarity 72.7%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 KSTNTFCCKPCC 17
Db 48 KSTGIFCKPSC 58

RESULT 3

hypothetical protein T23FL.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T23169
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: 219990

A:Accession: T23169
A:Status: preliminary; translated from GB/EMBL/DDB

A:Molecule type: DNA
A:Residues: 1-330 <WIL>
A:Cross-references: EMBL:281129; PIDN:CAB03405.1; GSPDB:GN00023; CESP:T23FL.6
C:Genetics:
A:Gene: CESP:T23FL.6
A:Map position: 5
A:Introns: 16/3
C:Superfamily: gliadin

Query Match 48.0%; Score 48; DB 2; Length 330;
Best Local Similarity 41.2%; Pred. No. 6.6;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CASELGSNTFCCKPCC 17
Db 80 CSQCCOSNTNTCCQPC 96

RESULT 4

hypothetical protein T15M6.10 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: B96614
R:Phelogiannis, A.; Ecker, J.R.; Palm, C.J.; Federpspiel, N.A.; Kaul, S.; White, O.; A.
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Mar
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tai
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: B96614
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STO>
A:Cross-references: GB:AE005173; NID:911067306; PIDN:AAG2883.1; GSPDB:GN00141
C:Genetics:
A:Gene: T15M6.10
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein F24M12.210

Query Match 45.0%; Score 45; DB 2; Length 371;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ASLGSNTFCCKPCC 17
Db 224 SKTEKSKRTFCVLPCC 239

RESULT 5

epidermal growth factor receptor homolog precursor - fluke (Schistosoma mansoni)

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Schistosoma mansoni
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A45558; S27836
R:Shoemaker, C.B.; Ramachandran, H.; Lande, A.; dos Reis, M.G.; Stein, L.D.
Mol. Biochem. Parasitol. 53, 17-32, 1992
A:Title: Alternative splicing of the Schistosoma mansoni gene encoding a homologue
A:Reference number: A45558; MUID:92365727
A:Accession: A45558
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-1717 <SHO>
A:Cross-references: EMBL:M86396; NID:g160957; PIDN:AA29866.1; PID:g160958
A:Note: sequence extracted from NCBI backbone (NCBIP:111129)
C:Genetics:
A:Gene: SER
C:Superfamily: fluke epidermal growth factor receptor homolog 1; protein kinase homolog
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; phosphoprotein
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-1717/Product: epidermal growth factor receptor homolog 1 #status predicted <MAT>
F:1018-1323/Domain: protein kinase homolog <KIN>
F:1026-1034/Region: protein kinase ATP-binding motif

Query Match 44.0%; Score 44; DB 1; Length 1717;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNPFCKRP 16
||:|||||
Db 1311 ELMRTNPFCKRP 1323

RESULT 6
D71882
hypothetical protein jhp0825 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
C:Accession: D71882
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120357
A:Accession: D71882
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-100 <ARN>
A:Cross-references: GB:AE001512; GB:AE001439; NID:g4155392; PIDN:AMD06402.1; PID:g415539
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0825
C:Superfamily: conserved hypothetical protein HI0711

Query Match 43.0%; Score 43; DB 2; Length 100;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 5 LGSSTNPFCKRP 16
||:|||||
Db 82 LGSSTNPFCKRP 93

RESULT 7
A39035
ribonuclease-related anti-tumor protein - northern leopard frog (fragment)
C:Species: Rana pipiens (northern leopard frog)
C:Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 30-Jun-1993
C:Accession: A39035
R:Ardelet, W.; Mikulski, S.M.; Shogen, K.
J. Biol. Chem. 266, 245-251, 1991
A:Title: Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and earl
A:Reference number: A39035; MUID:91093131
A:Accession: A39035
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <ARD>
C:Superfamily: pancreatic ribonuclease

Query Match 43.0%; Score 43; DB 2; Length 104;
Best Local Similarity 61.5%; Pred. No. 14;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 1 CASELCKSTNPFCK 13
|:|||||
Db 75 CKYKLRKSTNPFCK 87

RESULT 8
S64904
probable membrane protein YLR072w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L2321
C:Species: Saccharomyces cerevisiae
C:Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 05-Nov-1999
C:Accession: S64904
R:Pohl, T.M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64899
A:Accession: S64904
A:Molecule type: DNA
A:Residues: 1-693 <POH>
A:Cross-references: EMBL:Z73244; NID:g1360427; PID:e245792; PID:g1360428; GSPDB:GN
F:634-650/Domain: transmembrane #status predicted <TM>

Query Match 43.0%; Score 43; DB 2; Length 693;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 ASELCKSTNPFCKRP 15
|:|||||
Db 358 SSSLSSTTYICRP 371

RESULT 9
T08852
lustrin A - California red abalone
C:Species: Haliotis rufescens (California red abalone)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08852
R:Shen, X.; Belcher, A.M.; Hansma, P.K.; Stucky, G.D.; Morse, D.E.
J. Biol. Chem. 272, 32472-32481, 1997
A:Title: Molecular cloning and characterization of lustrin A, a matrix protein fr
A:Reference number: Z16496; MUID:98070424
A:Accession: T08852
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1428 <SHE>
A:Cross-references: EMBL:AF023459; NID:g2723361; PIDN:AAB95154.1; PID:g2723362
A:Experimental source: tissue type mantle (shell and pearl nacre); cell type pall
C:Superfamily: antileukoprotease repeat homolog
C:Keywords: extracellular matrix; extracellular protein
F:1382-1426/Domain: antileukoprotease repeat homology <ALP>

Query Match 42.0%; Score 42; DB 2; Length 1428;
Best Local Similarity 43.8%; Pred. No. 2e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 CASELCKSTNPFCKRP 16
|:|||||
Db 1334 CRTNLOCPSTNYCKRP 1349

RESULT 10
B38346
ultra-high-sulfur keratin 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

C:Accession: A38660; B38346
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogel, G.
J. Biol. Chem. 265, 4024, 1991
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38660; MUID:91154184
A:Accession: A38660
A:Molecule type: DNA
A:Residues: 1-223 <MO2>
A:Cross-references: GB:M37759; NID:g200963; PIDN:AAA40107.1; PID:g200962
A:Note: this is a correction
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogel, G.
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38346; MUID:91065960
A:Accession: B38346
A:Molecule type: DNA
A:Residues: 1-21, 'GGCGSGCGGCGSGNCGGCGSSCCCKPVC', 22-40, 'GSS', 44-45, 'G', 47-48, 'S', 50, 'GSS',
<MOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
A:Note: the sequence reported in this paper has been corrected. See A38660
C:Superfamily: ultra-high-sulfur keratin

Query Match
Best Local Similarity 41.0%; Score 41; DB 2; Length 223;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
OY 1 CASELGKSTNTFCCKPPC 17
DB 10 CGSGCGGCGGCGSSCCCKPVC 26

RESULT 11
A38346
ultra-high-sulfur keratin 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 24-Sep-1999
C:Accession: A38346
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogel, G.
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38346; MUID:91065960
A:Accession: A38346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <MOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
C:Superfamily: ultra-high-sulfur keratin

Query Match
Best Local Similarity 41.0%; Score 41; DB 2; Length 230;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
OY 1 CASELGKSTNTFCCKPPC 17
DB 31 CGSMCGGCGGSSCCCKPVC 47

RESULT 12
A84947
lybB protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: A84947
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A:Reference number: A84930; MUID:20445173
A:Accession: A84947
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: lytB; B0147
C:Superfamily: penicillin tolerance protein

Query Match
Best Local Similarity 41.0%; Score 41; DB 2; Length 319;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 3 SELGKSTNTFCCK 14
DB 232 AELGKETGFTK 243

RESULT 13
B45511
chitinase (EC 3.2.1.14) precursor, basic - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 22-Jun-1999
C:Accession: B45511
R:Samac, D.A.; Hironaka, C.M.; Yallaly, P.E.; Shah, D.M.
Plant Physiol. 93, 907-914, 1990
A:Title: Isolation and characterization of the genes encoding basic and acidic chitinases
A:Reference number: A45511
A:Accession: B45511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <SAM>
A:Cross-references: GB:M38240; NID:g166665; PIDN:AAA32769.1; PID:g166666
C:Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homolog
C:Keywords: glycosidase, hydrolase; polysaccharide degradation
F:34-76/Domain: hevein chitin-binding domain homology <HCB>
F:89-32/Domain: plant chitinase homology <PCH>

Query Match
Best Local Similarity 41.0%; Score 41; DB 2; Length 335;
Matches 9; Conservative 1; Mismatches 6; Indels 2; Gaps 1;
OY 1 CASELGKSTNT-FCCKP 16
DB 50 CSESEFGWCGNTEPYCKP 67

RESULT 14
AC3574
acetate kinase homolog AckA2 [imported] - Listeria innocua (strain C1p11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1574
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bio
D.; Domínguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entlian, K.D.; Fish
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, J.; Vazquez-Boland, J.A.; Voss, H.; Weh
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1574
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96363.1; PID:g16413591; GSPDB:GN00178
C:Genetics:
A:Gene: AckA2
C:Superfamily: acetate kinase

Query Match
Best Local Similarity 41.0%; Score 41; DB 2; Length 397;
Matches 72.7%; Pred. No. 95;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11
 11 | 111 | 11
 DB 214 CAIEAGKSVNT 224

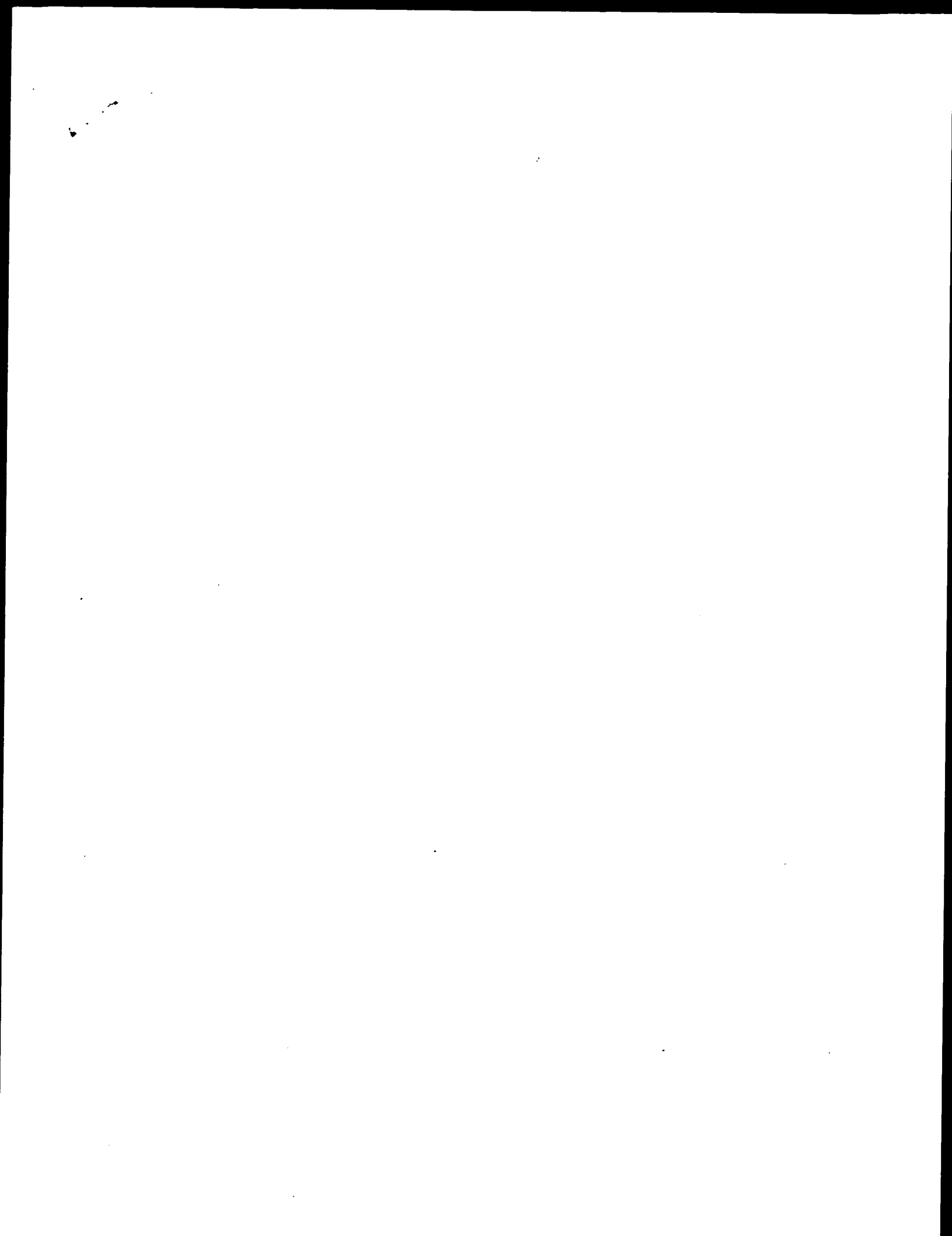
RESULT 15

AH1220
 acetate kinase homolog AckA2 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AH1220
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-397 <GIA>
 A:Cross-References: GB:NC_003210; PIDN:CAC9246.1; PID:gl6410584; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: AckA2
 C:Superfamily: acetate kinase

Query Match 41.0%; Score 41; DB 2; Length 397;
 Best Local Similarity 72.7%; Pred. No. 95;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CASELGKSTNT 11
 11 | 111 | 11
 DB 214 CAIEAGKSVNT 224

Search completed: June 17, 2002, 15:58:22
 Job time: 191 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:27 ; Search time 44.84 Seconds
(without alignments)
14.680 Million cell updates/sec

Title: US-09-761-636A-8
Perfect score: 100
Sequence: 1 CASELGRSTNTECKPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	80.0	354	1	VEGD_HUMAN
2	77	77.0	326	1	VEGD_RAT
3	77	77.0	358	1	VEGC_MOUSE
4	59	59.0	415	1	VEGC_MOUSE
5	58	58.0	419	1	VEGC_HUMAN
6	48	48.0	211	1	ADA_BACSU
7	46	46.0	905	1	YD83_HUMAN
8	46	46.0	1112	1	SUS1_HUMAN
9	43	43.0	104	1	RN30_RANPI
10	41	41.0	319	1	LYTB_BUCAI
11	41	41.0	322	1	CHIT_ARATH
12	41	41.0	642	1	UL77_HCVNA
13	41	41.0	685	1	CFAT_HCVNA
14	41	41.0	988	1	PKC1_SCHPO
15	41	41.0	1174	1	KPC1_COCH
16	41	41.0	1294	1	YOH5_YEAST
17	40.5	40.5	604	1	PGH2_CAVPO
18	40	40.0	224	1	Y662_MYCTU
19	40	40.0	328	1	CHIT_PHAVU
20	40	40.0	353	1	ADA_SALTY
21	40	40.0	500	1	RBUK_HUMAN
22	40	40.0	526	1	RBUK_MOUSE
23	40	40.0	594	1	CAFA_YERPE
24	40	40.0	833	1	IBB3_PHAVU
25	39	39.0	85	1	IBB3_PHAVU
26	39	39.0	85	1	IBB3_PHAVU
27	39	39.0	139	1	UTCH_HCVNA
28	39	39.0	322	1	CHIT2_BRANA
29	39	39.0	354	1	ADA_ECOLI
30	39	39.0	409	1	YNV5_CAEEL
31	39	39.0	758	1	PARC_RHIME
32	39	39.0	1151	1	KPC1_YEAST
33	39	39.0	1786	1	LMBI_HUMAN

34	38.5	38.5	230	1	UL16_HCVNA	P16757 human cytom
35	38	38.0	83	1	IBB3_SOYBN	P01064 glycine max
36	38	38.0	263	1	SGS3_DROYA	P13728 drosophila
37	38	38.0	471	1	VSIB_PTYBB	P07208 trypanosoma
38	38	38.0	982	1	HBL1_CABEL	Q99014 caenorhabdi
39	38	38.0	1139	1	KPC1_TIRE	Q99014 trichoderma
40	38	38.0	1142	1	KPC1_NUCR	P87253 neurospora
41	38	38.0	1370	1	2261_HUMAN	Q14202 homo sapien
42	38	38.0	2282	1	ZAN_RABIT	P57999 oryctolagus
43	37.5	37.5	681	1	MP10_HUMAN	000566 homo sapien
44	37	37.0	111	1	RL23_CHLMU	Q9P116 chlamydia m
45	37	37.0	111	1	RL23_CHLNP	Q9Z7G9 chlamydia p

ALIGNMENTS

RESULT 1
VEGD_HUMAN STANDARD: PRT: 354 AA.
ID 043915:
AC 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).
GN FIGF OR VEGFD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D";
RL Genomics 42:483-488(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98110120; PubMed=9479493;
RA Rocchianni M., Lestini M., Luddi A., Orlandini M., Franco B.,
RA Rossi E., Ballabio A., Zuffardi O., Oliviero S.;
RT "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
RT between the FIGA and the GRP genes";
RL Genomics 47:207-216(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98118549; PubMed=9435229;
RA Achen M.G., Veltsch M., Kuk E., Meekinen T., Vitall A., Wilks A.F.,
RA Allitalo K., Stacker S.A.;
RT "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
RT tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
RN [4]
RP PROCESSING, AND SEQUENCE OF 89-94: 100-105 AND 206-213.
RX MEDLINE=20011413; PubMed=10542248;
RA Stacker S.A., Stenvers K.L., Caesar C., Vitall A., Domagala T.,
RA Nice E.C., Roufai S., Simpson R.J., Moritz R., Karpanen T.,
RA Allitalo K., Achen M.G.;
RT "Biosynthesis of vascular endothelial growth factor-D involves
RT proteolytic processing which generates non-covalent homodimers";
RL J. Biol. Chem. 274:32127-32136(1999).
RN [5]
RP FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
RN and endothelial cell growth, stimulating their proliferation and
RN migration and also has effects on the permeability of blood
RN vessels. May function in the formation of the venous and lymphatic
RN vascular systems during embryogenesis, and also in the maintenance
RN of differentiated lymphatic endothelium in adults. Binds and
RN activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
RN -i- SUBUNIT: Homodimer; non-covalent and antiparallel.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Highly expressed in lung, heart, small
 CC intestine and fetal lung, and at lower levels in skeletal muscle,
 CC colon, and pancreas.
 CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
 CC linked by disulfide bonds before secretion. The fully processed
 CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 CC bound by non-covalent interactions.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC
 CC EMBL: D86630; BAA24264.1; -
 CC EMBL: Y12863; CAA73370.1; -
 CC EMBL: Y12864; CAA73371.1; -
 CC EMBL: Y12865; CAA73371.1; JOINED.
 CC EMBL: Y12866; CAA73371.1; JOINED.
 CC EMBL: Y12867; CAA73371.1; JOINED.
 CC EMBL: Y12868; CAA73371.1; JOINED.
 CC EMBL: Y12869; CAA73371.1; JOINED.
 CC EMBL: Y12870; CAA73371.1; JOINED.
 CC EMBL: AJ000185; CAA03942.1; -
 CC MIM: 300091; -
 CC HSSP: P15692; 1VPP.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF_1.
 CC ProDom: PD001629; PDGF_1.
 CC SMART: SM00141; PDGF_1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS0278; PDGF_2; 1.
 CC KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
 CC Cleavage on pair of basic residues; Multigene family.
 CC FT SIGNAL 1 21
 CC FT PROPEP 88
 CC FT CHAIN 205
 CC FT PROPEP 354
 CC FT DOMAIN 222 318
 CC
 CC FT REPEAT 222 237
 CC FT REPEAT 258 273
 CC FT REPEAT 277 293
 CC FT REPEAT 301 318
 CC FT DISULFID 111 153
 CC FT DISULFID 142 189
 CC FT DISULFID 146 191
 CC FT DISULFID 136 136
 CC FT DISULFID 145 145
 CC FT CARBOHYD 155 155
 CC FT CARBOHYD 185 185
 CC FT CARBOHYD 287 287
 CC SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;
 CC
 CC Query Match 80.0%; Score 80; DB 1; Length 354;
 CC Best Local Similarity 93.8%; Pred. No. 6; Se-06;
 CC Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AC 035251;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
 DE growth factor) (FIGF).
 GN FIGF OR VEGFD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RC Yamada Y., Hirata Y., Nezu J., Shitane M.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
 CC and endothelial cell growth, stimulating their proliferation and
 CC migration and also has effects on the permeability of blood
 CC vessels. May function in the formation of the venous and lymphatic
 CC vascular systems during embryogenesis, and also in the maintenance
 CC of differentiated lymphatic endothelium in adults. Binds and
 CC activates VEGFR-3 (Flt4) receptor (By similarity).
 CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
 CC linked by disulfide bonds before secretion. The fully processed
 CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 CC bound by non-covalent interactions (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC
 CC EMBL: AF014827; AAB66557.1; -
 CC HSSP: P15692; 1VPP.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF_1.
 CC ProDom: PD001629; PDGF_1.
 CC SMART: SM00141; PDGF_1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS0278; PDGF_2; 1.
 CC KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
 CC Cleavage on pair of basic residues; Multigene family.
 CC FT SIGNAL 1 21
 CC FT PROPEP 93
 CC FT CHAIN 210
 CC FT PROPEP 326
 CC FT DOMAIN 227 317
 CC
 CC FT REPEAT 227 242
 CC FT REPEAT 263 278
 CC FT REPEAT 282 298
 CC FT REPEAT 306 317
 CC FT DISULFID 116 158
 CC FT DISULFID 147 194
 CC FT DISULFID 151 196
 CC FT DISULFID 141 141
 CC FT DISULFID 150 150
 CC FT CARBOHYD 160 160
 CC FT CARBOHYD 190 190
 CC FT CARBOHYD 292 292
 CC SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;
 CC
 CC Query Match 77.0%; Score 77; DB 1; Length 326;

Best Local Similarity 87.5%; Pred. No. 2e-05; Mismatches 1; Indels 0; Gaps 0;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ASELGKSTNTECKPCC 17
111111:1111111111

DB 126 ASELGKSTNTECKPCC 141

RESULT 3
VEGC_MOUSE STANDARD: PRT: 358 AA.

AC P97946;
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (Figf).
DE FIGF OR VEGF.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Fibroblast;
RX MEDLINE=97030254; PubMed=8676195;
RA Orlandini M., Marconcini L., Ferruzzi R., Oliviero S.;
RT "Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family";
RT Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=97349118; PubMed=9205122;
RX Yanada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor, VEGF-D";
RT Genomics 42:483-488(1997).
RL [3]
RN DEVELOPMENTAL STAGE.
RP MEDLINE=98288130; PubMed=9622638;
RX Avantiagato V., Orlandini M., Acampora D., Oliviero S., Simeone A.;
RT "Avascular expression pattern of the murine figf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family";
RT Mech. Dev. 73:221-224(1998).
RL [4]
RN RECEPTOR SPECIFICITY.
RP MEDLINE=21276411; PubMed=11279005;
RX Baldwin M.E., Catimel B., Nice E.C., Roufail S., Hall N.E.,
RA Stenvers K.L., Karkkainen M.J., Allitalo K., Stacker S.A., Achen M.G.;
RT "The specificity of receptor binding by vascular endothelial growth factor-d is different in mouse and man.";
RL J. Biol. Chem. 276:19166-19171(2001).
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis, and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor.
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.
CC -1- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several body structures and organs of the embryo such as limb buds, acoustic ganglion, teeth, heart, anterior pituitary as well as lung and kidney mesenchyme, liver, derma, and peritoneum of the vertebral column.
CC -1- INDUCTION: By the transcription factor c-fos.
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward

CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS) bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X09572; CAA67892.1; -;
CC EMBL: D89628; BAA14002.1; -;
CC HSSP: P15692; 1VPF.
CC MGD: MGI:108037; Figf.
CC InterPro: IPR000072; PDGF.
CC Pfam: PF00341; PDGF; 1.
CC ProDom: PD001629; PDGF; 1.
CC SMART: SM00141; PDGF; 1.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS50278; PDGF_2; 1.
CC MtoGen: Growth factor; Glycoprotein; Signal; Repeat;
CC Cleavage on pair of basic residues; Multigene family.
CC SIGNAL 1 21
CC PROPEP 92 93
CC CHAIN 94 210
CC PROPEP 211 358
CC DOMAIN 227 323
CC REPEAT 227 242
CC REPEAT 263 278
CC REPEAT 282 298
CC REPEAT 306 323
CC DISULFID 116 158
CC DISULFID 147 194
CC DISULFID 151 196
CC DISULFID 141 141
CC DISULFID 150 150
CC CARBOHYD 160 160
CC CARBOHYD 190 190
CC CARBOHYD 292 292
CC SEQUENCE 358 AA; 40908 MW; 6636B17BF07037C CRC64;
Query Match Score 77.0%; DB 1; Length 358;
Best Local Similarity 87.5%; Pred. No. 2e-05; Mismatches 1; Indels 0; Gaps 0;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ASELGKSTNTECKPCC 17
111111:1111111111

DB 126 ASELGKSTNTECKPCC 141

RESULT 4
VEGC_MOUSE STANDARD: PRT: 415 AA.

AC P97953;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor C precursor (VEGF-C) (Vascular endothelial growth factor related protein) (VRF) (Flt4 ligand) (Flt4-l).
DE FIGF.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.

CC STRAIN-BALB/C;
 CC MEDLINE-97164697; PubMed-9012504;
 RA Kulk E., Lybousaki A., Taira S., Kaipainen A., Jeltsch M.,
 RA Joukov V., Altalo K.;
 RT "VEGF-C receptor binding and pattern of expression with VEGFR-3
 RT suggests a role in lymphatic vascular development.";
 RN Development 122:3829-3837(1996).
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-126.
 RC STRAIN-BALB/C;
 RA MEDLINE-97388482; PubMed-9247316;
 RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
 RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charette A.,
 RA Giannotti J., Finerly H., Zollner R., Beler D.R., Leak L.V.,
 RA Turner K.J., Wood C.R.;
 RT "Characterization of murine Flt4 ligand/VEGF-C";
 RT Oncogene 15:613-618(1997).
 CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
 CC cell growth, stimulating their proliferation and migration and
 CC also has effects on the permeability of blood vessels. May
 CC function in angiogenesis of the venous and lymphatic vascular
 CC systems during embryogenesis, and also in the maintenance of
 CC differentiated lymphatic endothelium in adults. Binds and
 CC activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
 CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expression detected in mesenchymal cells of
 CC postimplantation embryos, particularly in the regions where the
 CC lymphatic vessels sprout from embryonic veins, such as
 CC the perimetaphic, axillary and jugular regions, and in the
 CC developing mesenterium. Expressed in adult heart, brain, spleen,
 CC lung, liver, skeletal muscle and kidney.
 CC -1- PM: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3, but only the fully processed form could activate VEGFR-2.
 CC bonds. Before secretion, a cleavage occurs between arg-227 and
 CC ser-228 producing a heterotrimer. The next extracellular step
 CC of the processing removes the N-terminal propeptide. Finally the
 CC mature VEGF-C is composed mostly of two VEGF homology domains
 CC (VHDS) bound by non-covalent interactions (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, obstruction
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U73620; AAC52984.1; -;
 DR EMBL: U58112; AAB46707.1; -;
 DR HSSP: P15692; IVPF.
 DR MGD: MGI:109124; Vegfc.
 DR InterPro: IPR004153; CXKC.
 DR InterPro: IPR002400; GF_CysKnot.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF03128; CXKC; 5.
 DR Pfam: PF00341; PDGF; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR ProDom: PD001629; PDGF; 1.
 DR SMART: SM00141; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1;
 DR PROSITE: PS00278; PDGF_2; 1.
 KW Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
 KW Cleavage on pair of basic residues; Multigene family;
 FT SIGNAL 1 31
 FT PROPEP 32 107
 FT CHAIN 108 223
 FT PROPEP 224 415
 FT DOMAIN 276 358
 X(1,3)-C.

FT REPEAT 276 291
 FT REPEAT 300 315
 FT REPEAT 324 339
 FT REPEAT 343 358
 FT DISULFID 127 159
 FT DISULFID 158 205
 FT DISULFID 162 207
 FT DISULFID 152 152
 FT DISULFID 151 151
 FT CARBOHYD 171 171
 FT CARBOHYD 201 201
 FT CARBOHYD 236 236
 SQ SEQUENCE 415 AA; 46471 MW; D903DB3CECC659D6 CRC64;
 Query Match 59.0%; Score 59; DB 1; Length 415;
 Best Local Similarity 71.4%; Pred. No. 0.023;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Oy 4 ELGKSTNPFCKRPPC 17
 Db 139 EFGAATNPFCKRPPC 152
 RESULT 5
 VEGC_HUMAN
 ID VEGC_HUMAN STANDARD: PRT; 419 AA.
 AC PA9767;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor C precursor (VEGF-C) (Vascular
 DE endothelial growth factor related protein) (VRF) (Flt4 ligand) (Flt4-
 DE L).
 GN VEGFC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
 RA MEDLINE-96178224; PubMed-8617204;
 RA Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kulk E.,
 RA Saksela O., Kalkkinen N., Altalo K.;
 RT "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
 RT the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases";
 RL EMBO J. 15:290-298(1996).
 RN [2]
 RP ERRATUM.
 RX MEDLINE-96203094; PubMed-8612600;
 RA Joukov V., Pajusola K., Kaipainen A., Chillov D., Lahtinen I., Kulk E.,
 RA Saksela O., Kalkkinen N., Altalo K.;
 RL EMBO J. 15:1751-1751(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Specific tumor;
 RX MEDLINE-96312526; PubMed-8700872;
 RA Lee J., Gray A., Yuan J., Luo S.-M., Avraham H., Wood W.I.;
 RT "Vascular endothelial growth factor-related protein: a ligand and
 RT specific activator of the tyrosine kinase receptor Flt4";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97388482; PubMed-9247316;
 RA Fitz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
 RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charette A.,
 RA Giannotti J., Finerly H., Zollner R., Beler D.R., Leak L.V.,
 RA Turner K.J., Wood C.R.;
 RT "Characterization of murine Flt4 ligand/VEGF-C";
 RT Oncogene 15:613-618(1997).
 RN [5]
 RP SEQUENCE OF 32-41; 112-121 AND 228-233, AND MUTAGENESIS OF ARG-227.
 RX MEDLINE-97377029; PubMed-9233800;

RA Jéoukov V., Sorsa T., Kumar V., Jeltsch M., Claesson-Welsh L., Cao Y.,
 RA Sakela O., Kalkkinen N., Alltalo K.;
 RT "Proteolytic processing regulates receptor specificity and activity of
 RT VEGF-C".
 RL EMBL J. 16:3898-3911(1997).
 CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
 CC cell growth, stimulating their proliferation and migration and
 CC also has effects on the permeability of blood vessels. May
 CC function in angiogenesis of the venous and lymphatic vascular
 CC systems during embryogenesis, and also in the maintenance of
 CC differentiated lymphatic endothelium in adults. Binds and
 CC activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
 CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Spleen, lymph node, thymus, appendix, bone
 CC marrow, heart, placenta, ovary, skeletal muscle, prostate, testis,
 CC colon and small intestine and fetal liver, lung and kidney, but
 CC not in peripheral blood lymphocyte.
 CC -1- PFM: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3, but only the fully processed form could activate VEGFR-2.
 CC VEGF-C first form an antiparallel homodimer linked by disulfide
 CC bonds. Before secretion, a cleavage occurs between arg-227 and
 CC ser-228 producing an heterotetramer. The next extracellular step
 CC of the processing removes the N-terminal propeptide. Finally the
 CC mature VEGF-C is composed mostly of two VEGF homology domains
 CC (VHDS) bound by non-covalent interactions.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X94216; CAA63907.1; -
 DR EMBL: U43142; AA85214.1; -
 DR EMBL: U58111; AA802909.1; -
 DR HSSP: P15692; 1VPF.
 DR MIM: 601528; -
 DR InterPro: IPR004153; GCXCXC.
 DR InterPro: IPR002400; GF_Cysknot.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF003128; GCXCXC; 5.
 DR Pfam: PF003411; PDGF; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PRODOM: PD001629; PDGF; 1.
 DR SMART: SM00141; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
 KW Cleavage on pair of basic residues; Multigene family.
 FT SIGNAL 1 31
 FT PROPEP 32 111
 FT CHAIN 112 227
 FT PROPEP 228 419
 FT DOMAIN 280 362
 FT REPEAT 280 295
 FT REPEAT 304 319
 FT REPEAT 328 343
 FT REPEAT 347 362
 FT DISULFID 131 173
 FT DISULFID 162 209
 FT DISULFID 166 211
 FT DISULFID 156 156
 FT DISULFID 165 165
 FT DISULFID 175 175
 FT CARBOHYD 205 205
 FT CARBOHYD 240 240
 FT MUTAGEN 227 227

OR 102.
 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 4 X 16 AA REPEATS OF C-X(10)-C-X-C-
 X(1,3)-C.
 1.
 2.
 3.
 4.
 INTRACHAIN (BY SIMILARITY).
 INTRACHAIN (BY SIMILARITY).
 INTRACHAIN (BY SIMILARITY).
 INTRACHAIN (BY SIMILARITY).
 INTERCHAIN (BY SIMILARITY).
 INTERCHAIN (BY SIMILARITY).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 R->S: NO PROTEOLYTIC PROCESSING AND LOWER

FT SQ SEQUENCE 419 AA; 46883 MM; 9F598719DB3E014F CRC64; EFFECT ON VEGFR-2 AND VEGFR-3.
 Query Match 58.0%; Score 58; DB 1; Length 419;
 Best Local Similarity 71.4%; Pred. No. 0.034;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 ELKSTNFFCRPPC 17
 1 -11111111
 Db 143 EFGVATNFFCRPPC 156
 RESULT 6
 ADAA_BACSU STANDARD; PRT; 211 AA.
 ID ADAA_BACSU
 AC P19219;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Methylphosphotriester-DNA alkyltransferase.
 GN ADAA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=91016831; PubMed=2120677;
 RA Morohoshi F., Hayashi K., Munakata N.;
 RT "Bacillus subtilis ada operon encodes two DNA alkyltransferases.";
 RL Nucleic Acids Res. 18:5473-5480(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
 RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
 RL Bacillus subtilis chromosome.";
 RN Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP MUTAGENESIS.
 RX MEDLINE=92078089; PubMed=1744039;
 RA Morohoshi F., Hayashi K., Munakata N.;
 RT "Molecular analysis of Bacillus subtilis ada mutants deficient in the
 RT adaptive response to simple alkylating agents.";
 RL J. Bacteriol. 173:7834-7840(1991).
 CC -1- FUNCTION: METHYLPHOSPHOTRIESTER-DNA ALKYLTRANSFERASE IS ONE OF
 CC TWO GENES REQUIRED FOR THE ADAPTIVE RESPONSE OF BACTERIA TO
 CC ALKYLATING AGENTS. IT ACCEPTS A METHYL GROUP FROM METHYLPHOSPHO-
 CC TRIESTERS AND THEN ACTS AS A TRANSCRIPTIONAL ACTIVATOR OF THE
 CC ADA OPERON.
 CC -1- SIMILARITY: TO THE N-TERMINAL REGION OF E. COLI ADA PROTEIN AND
 CC THE C-TERMINAL REGION OF THE ARAC/XYLX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X53399; CAA37475.1; -
 DR EMBL: AB006424; BAA33074.1; -
 DR EMBL: Z99104; CAB11957.1; -
 DR EMBL: Z99105; CAB11974.1; -
 DR PIR: S11483; XUBSM.
 DR HSSP: P06134; 1ADN.
 DR Subtilist: BG10166; adaA.
 DR InterPro: IPR004026; Ada_Zn_bind.
 DR InterPro: IPR000005; HTHARAC.
 DR Pfam: PF02805; Ada_Zn_binding; 1.

DR Pfam: PF00165; HTH_Arac; 1.
 DR PRINTS; PRO0032; HTHARAC.
 DR SMART; SM00342; HTH_ARAC; 1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
 KW Transferrase; Methyltransferase; transcription regulation; Activator;
 KM DNA-binding; DNA repair; Zinc; Metal-binding; Complete proteome.
 FT METAL 54
 FT METAL 54
 FT METAL 58
 FT METAL 85
 FT METAL 88
 FT METAL 118
 FT ACT_SITE 137
 FT ACT_SITE 85
 FT MUTAGEN 87
 FT SEQUENCE 211 AA; 24299 MW; 920931082527EC27 CRC64;

Query Match 48.0%; Score 48; DB 1; Length 211;
 Best Local Similarity 72.7%; Pred. No. 0.76;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 KSTNFCRPPC 17
 DB 48 KSTGIFCKPSC 58

RESULT 7
 YD83_HUMAN STANDARD; PRT; 905 AA.
 ID YD83_HUMAN
 AC Q9P2G4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein KIAA1383.
 GN KIAA1383
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP COVILLE G.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AB037804; BAA92621.1; ALT_INIT.
 DR EMBL; AL451083; CAC36343.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 905 AA; 100344 MW; B02FBD0EDAD78491 CRC64;

Query Match 46.0%; Score 46; DB 1; Length 905;
 Best Local Similarity 57.1%; Pred. No. 6.6;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 OY 3 SELGKSTNFCRPP 16
 : 11 11 11

DB 290 TELIDMETNFCRPP 303

RESULT 8
 SUS1_HUMAN STANDARD; PRT; 1112 AA.
 ID SUS1_HUMAN
 AC Q9GZRI; Q9JUV5; O94891;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE SUMO-1-specific protease 1 (EC 3.4.22.-) (Sentrin-specific protease
 DE SENP6) (Protease FKS66)
 GN SUSP1 OR SSP1 OR KIAA0797.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20261527; PubMed=10799485;
 RA Kim K.-I., Baek S.-H., Jeon Y.-D., Nishimori S., Suzuki T., Uchida S.,
 RA Shimbara N., Saitoh H., Tanaka K., Chung C.H.;
 RT "A new SUMO-1-specific protease, SUSP1, that is highly expressed in
 RT reproductive organs.";
 RL J. Biol. Chem. 275:14102-14106(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RX MEDLINE=20267842; PubMed=10806345;
 RA Yeh E.-T., Gong L., Kamitani T.;
 RT "Ubiquitin-like proteins: new wines in new bottles.";
 RL Gene 248:1-14(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Wang Y.-G.;
 RT "Identification of FKS66, a novel protein with protease activity.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 29-1112 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99087487; PubMed=9872452;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 5:277-286(1998).
 RL [1]
 RN FUNCTION: Protease that release SUMO-1 from its precursor
 CC sequence.
 CC SUBCELLULAR LOCATION: Cytoplasmic.
 CC TISSUE SPECIFICITY: Highly expressed in reproductive organs, such
 CC as testis, ovary, and prostate.
 CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY C48.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF196304; AAF04852.1; -;
 DR EMBL; AF307849; AAG29831.1; -;
 DR EMBL; AF306508; AAG30253.1; -;
 DR EMBL; AB018340; BAA34517.1; -;
 DR MIM: 605003; -;
 DR MEROPS: C48.004; -;
 DR InterPro: IPR003653; SUMO_protease.
 DR Hydrolase: Ubiquitin protease; Ubiquitin conjugation.
 KW DOMAIN 666 1112
 FT ACT_SITE 765 765
 FT BY SIMILARITY.

FT AGT_SITE 917 917 BY SIMILARITY.
 FT ACT_SITE 1030 1030 BY SIMILARITY.
 FT CONFLICT 121 121 T -> M (IN REF. 1 AND 4).
 FT CONFLICT 1043 1043 Q -> E (IN REF. 1 AND 4).
 FT CONFLICT 1106 1106 Y -> C (IN REF. 1 AND 4).
 FT CONFLICT 1112 1112 AA: 126144 MW: B6F53875C383A46 CRC64;
 SQ SEQUENCE

Query Match 46.0%; Score 46; DB 1; Length 1112;
 Best Local Similarity 47.4%; Pred. No. 8.1;
 Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 1 CASELCK--STNTECKPPC 17
 1:11:11 1 11 11
 Db 954 CSSEICGWHLKPTICKOPC 972

RESULT 9
 RN30_RANPI STANDARD; PRT: 104 AA.
 ID RN30_RANPI
 AC P22069;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE P-30 protein (EC 3.1.27.-) (Onconase).
 OS Rana pipiens (Northern leopard frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 OX NCBI_TaxID=8404;
 RN (1)
 RP SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=9103131; PubMed=1985896;
 RA Ardelt W., Mikulski S.M., Shogen K.;
 RT "Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases.";
 RL J. Biol. Chem. 266:245-251(1991).
 [2]
 RN (2)
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=9106156; PubMed=1438177;
 RA Mosimann S.C., Ardelt W., James M.N.G.;
 RT "Comparative molecular modeling and crystallization of P-30 protein: a novel antitumor protein of Rana pipiens oocytes and early embryos.";
 RL Proteins 14:392-400(1992).
 [3]
 RN (3)
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RX MEDLINE=94166079; PubMed=8120892;
 RA Mosimann S.C., Ardelt W., James M.N.G.;
 RT "Refined 1.7 A X-ray crystallographic structure of P-30 protein, an amphibian ribonuclease with anti-tumor activity.";
 RL J. Mol. Biol. 236:1141-1153(1994).
 CC -1- FUNCTION: BASIC PROTEIN WITH ANTIPROLIFERATIVE/CYTOTOXIC ACTIVITY AGAINST SEVERAL TUMOR CELL LINES IN VITRO. AS WELL AS ANTITUMOR IN VIVO. IT EXHIBITS A RIBONUCLEASE-LIKE ACTIVITY AGAINST HIGH MOLECULAR WEIGHT RIBOSOMAL RNA.
 CC -1- DEVELOPMENTAL STAGE: EARLY EMBRYOS (UP TO FOUR BLASTOMERE STAGE).
 CC -1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
 DR PDB: 1ONC; 31-JAN-94.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA; 1.
 DR Pfam: PF000535; RNaseA; 1.
 DR SMART: SM00092; RNase_Pc; 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC; 1.
 KW Hydroxylase; Nuclease; Endonuclease; 3D-structure.
 KM MOD_RES 1
 FT ACT_SITE 10 10 PYRROLIDONE CARBOXYLIC ACID.
 FT ACT_SITE 31 31
 FT ACT_SITE 97 97
 FT DISULFID 19 68
 FT DISULFID 30 75
 FT DISULFID 48 90

FT DISULFID 87 104
 FT HELIX 3 10
 FT STRAND 11 12
 FT HELIX 19 22
 FT TURN 23 24
 FT TURN 26 30
 FT STRAND 33 38
 FT STRAND 41 48
 FT TURN 49 50
 FT STRAND 55 58
 FT STRAND 63 70
 FT TURN 74 75
 FT STRAND 77 84
 FT STRAND 86 91
 FT TURN 92 93
 FT STRAND 94 101
 SQ SEQUENCE 104 AA: 11845 MW: 22A753C2F9E56B4 CRC64;

Query Match 43.0%; Score 43; DB 1; Length 104;
 Best Local Similarity 61.5%; Pred. No. 2.6;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CASELCKSTNTECK 13
 1:111111 1
 Db 75 CKYKLRKSTNTECK 87

RESULT 10
 LYT8_BUCAT STANDARD; PRT: 319 AA.
 ID LYT8_BUCAT
 AC P57247;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE LYT8 protein homolog.
 GN LYT8 OR BU147
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=TOXKO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- PATHWAY: Deoxyxylulose-5-phosphate pathway (DXP) of isoprenoid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE LYT8 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP001118; BAB12865.1; -
 DR InterPro: IPR003451; LYT8.
 DR Pfam: PF02401; LYT8; 1.
 KW Isoprene biosynthesis; Complete proteome.
 SQ SEQUENCE 319 AA: 35741 MW: 69B8AFC12DD09B8 CRC64;

Query Match 41.0%; Score 41; DB 1; Length 319;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 SELGKSTNTECK 14

```

Db      232 AELKGTGTFTK 243
RESULT 11
CHIT_ARATH STANDARD: PRT: 322 AA.
AC P19171;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Basic endochitinase precursor (EC 3.2.1.14).
GN ATG12500 OR T2E2.18 OR T2E22.119 OR MQC3.32 OR MQC3.34.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Sanac D.A., Hironaka C.M., Yallaly P.E., Shah D.M.;
RT Isolation and characterization of the genes encoding basic and
RL acidic chitinase in Arabidopsis thaliana.
RN [2]
RP Plant Physiol. 93:907-914(1990).
RC SEQUENCE FROM N.A.
RA MEDLINE=20363099; PubMed=10907853;
RT "Structural analysis of Arabidopsis thaliana chromosome 3 II.
RT Sequence features of the 4,251,695 bp regions covered by 90 PL, TAC
RT and BAC clones."
RL DNA Res. 7:217-221(2000).
RN [1]
RP FUNCTION: THIS PROTEIN FUNCTIONS AS A DEFENSE AGAINST CHITIN
CC CONTAINING FUNGAL PATHOGENS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-1-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -1- SUBCELLULAR LOCATION: VACUOLAR AND PROTOPLASTS
CC -1- TISSUE SPECIFICITY: HIGH CONSTRUCTIVE LEVEL IN ROOTS WITH LOWER
CC LEVELS IN LEAVES AND FLOWERING SHOOTS.
CC -1- INDUCTION: ETHYLENE INDUCES HIGH LEVELS OF SYSTEMIC EXPRESSION
CC -1- OF BASIC CHITINASE WITH EXPRESSION INCREASING WITH PLANT AGE.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS IA WHICH IS SIMILAR TO
CC CLASS IB IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL
CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M38240; AAA32769.1; ALT_INIT.
DR HSSP: P23951; 2BA.
DR InterPro: IPR001002; Chitin_bind.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; chitin_binding; 1.
DR Pfam: PF00182; Glyco_hydro_19; 1.
DR PRINTS: PR00451; CHITINBINDING.
DR ProDom: PD000574; Glyco_hydro_19; 1.
DR ProDom: PD000609; Chitin_bind; 1.
DR SMART: SM00270; ChitBD1; 1.
DR PROSITE: PS00026; CHITIN_BINDING; 1.
DR PROSITE: PS00773; CHITINASE_19_1; 1.
DR PROSITE: PS00774; CHITINASE_19_2; 1.
DR Hydrolase: Glycosidase; Chitin degradation; Chitin-binding; Signal.
FT SIGNAL 1
FT CHAIN 21 315 BASIC ENDOCHITINASE.

```

```

FT PROPEP 316 322 REMOVED IN MATURE FORM (PROBABLE).
FT DOMAIN 21 63 CHITIN-BINDING (BY SIMILARITY).
FT DISULFID 23 38 BY SIMILARITY.
FT DISULFID 32 44 BY SIMILARITY.
FT DISULFID 37 51 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
SQ SEQUENCE 322 AA; 34609 MW; C9AFEEAC544FCDD7 CRC64;

Query Match 41.0%; Score 41; DB 1; Length 322;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 9; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

QY 1 CASELGKSTNT--ECKP 16
Db 37 CSSEFGMCGNTEPYCKP 54

RESULT 12
UL77_HCMVA STANDARD: PRT: 642 AA.
AC P16726;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 01-DEC-1992 (Rel. 24, Last annotation update)
DE Virion protein UL77.
GN UL77.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Hornslell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Predelli E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
RN [1]
RP FUNCTION: VIRION PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL25,
CC EHV-1 36, EBV BHRF1, HCMV UL77, ILTV ORF2, AND VZV 34.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17403; CAA35392.1; -
DR PIR: S09840; Q08B5.
DR InterPro: IPR002493; UL25.
DR Pfam: PF01499; UL25; 1.
DR CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 642 AA; 71186 MW; A21EC24A95F460D5 CRC64;

Query Match 41.0%; Score 41; DB 1; Length 642;
Best Local Similarity 41.2%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 CASELGKSTNTCKPCKP 17
Db 94 CELDGGPSPRTLLQPPC 110

RESULT 13
CPAH_BOVIN

```

ID C3AH_BOVIN STANDARD: PRT: 685 AA.
AC 028085;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Complement factor H (H factor 1) (Fragments).
GN HFL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 17-685 FROM N.A., AND SEQUENCE OF 1-16.
RC TISSUE=Liver;
RX MEDLINE=96202005; PubMed=8615824;
RA Soames C.J., Day A.J., Slim R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
CC -1- FUNCTION: Factor H functions as a cofactor in the inactivation of
C3b by factor I and also increases the rate of dissociation of the
C3bb complex (C3 convertase) and the (C3b)NMB complex (C5
convertase) in the alternative pathway (By similarity).
CC -1- SIMILARITY: CONTAINS AT LEAST 13 SUSHI (SCR) DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: X98697; CAA67257.1; -
CC HSSP: P10998; IYVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi. 11.
DR SMART: SM00032; CCP. 11.
KW Complement alternate pathway; Plasma; Repeat; Sushi.
FT NON_TER 1 1
FT NON_CONS 16 17
FT DOMAIN <17 67 SUSHI 2.
FT DOMAIN 70 131 SUSHI 3.
FT DOMAIN 134 188 SUSHI 4.
FT DOMAIN 191 246 SUSHI 5.
FT DOMAIN 249 309 SUSHI 6.
FT DOMAIN 311 366 SUSHI 7.
FT DOMAIN 370 429 SUSHI 8.
FT DOMAIN 431 488 SUSHI 9.
FT DOMAIN 491 547 SUSHI 10.
FT DOMAIN 553 609 SUSHI 11.
FT DOMAIN 614 668 SUSHI 12.
FT DOMAIN 675 >685 SUSHI 13.
FT DISULFID 39 66 BY SIMILARITY.
FT DISULFID 71 117 BY SIMILARITY.
FT DISULFID 103 130 BY SIMILARITY.
FT DISULFID 135 176 BY SIMILARITY.
FT DISULFID 162 187 BY SIMILARITY.
FT DISULFID 192 234 BY SIMILARITY.
FT DISULFID 219 245 BY SIMILARITY.
FT DISULFID 250 297 BY SIMILARITY.
FT DISULFID 280 308 BY SIMILARITY.
FT DISULFID 312 354 BY SIMILARITY.
FT DISULFID 339 365 BY SIMILARITY.
FT DISULFID 371 417 BY SIMILARITY.
FT DISULFID 400 428 BY SIMILARITY.
FT DISULFID 432 476 BY SIMILARITY.
FT DISULFID 459 487 BY SIMILARITY.
FT DISULFID 492 534 BY SIMILARITY.
FT DISULFID 520 546 BY SIMILARITY.
FT DISULFID 554 597 BY SIMILARITY.
FT DISULFID 583 608 BY SIMILARITY.

FT DISULFID 615 656 BY SIMILARITY.
FT DISULFID 642 667 BY SIMILARITY.
FT NON_TER 685 685
SQ SEQUENCE 685 AA; 77536 MW; 69FC9DC8D50E872 CRC64;
Query Match 41.0%; Score 41; DB 1; Length 685;
Best Local Similarity 46.7%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Oy 2 ASSEKSTNTFCXP 16
Db 124 SAETPKCEIFCKXP 138
RESULT 14
PK1_SCHPO STANDARD: PRT: 988 AA.
ID PK1_SCHPO
AC P36582;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Protein kinase C-like 1 (EC 2.7.1.1)
PK1 OR SPAC22H10.01C OR SPAC1768.14C.
OS Schizosaccharomyces pombe (fission yeast).
GN Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetes; Schizosaccharomycetales;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93259141; PubMed=8491190;
RA Toda T., Shimanuki M., Yanagida M.;
RT "Two novel protein kinase C-related genes of fission yeast are
essential for cell viability and implicated in cell shape control."
EMBO J. 12:1987-1995(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE CONTROL OF THE CELL SHAPE. TARGET
OF THE INHIBITOR STRAUSSPORINE.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: D14337; BA03267.1; -
CC EMBL: Z69730; CAA93602.1; -
CC EMBL: Z69795; CAA93697.1; -
CC PIR: S35362; S35362.
CC HSSP: O63450; 1A06.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000961; pkinase_C.
DR InterPro: IPR000861; REM_repeat.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF02185; Hrl; 2.
DR Pfam: PF00069; pkinase_1.
DR Pfam: PF00433; pkinase_C; 1.
DR SMART: SM00109; C1; 2.
DR SMART: SM00239; C2; 1.

DR SMART; SM00074; HRL; 2.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; zinc;
 Phorbol-ester binding; Repeat.
 FT DOMAIN 414 461
 FT DOMAIN 481 530 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 664 923 PHORBOL-ESTER AND DAG BINDING 2.
 FT NP_BIND 670 678 PROTEIN KINASE.
 FT BINDING 693 693 ATP (BY SIMILARITY).
 FT ACT_SITE 789 789 ATP (BY SIMILARITY).
 FT CONFLICT 27 35 BY SIMILARITY.
 SQ SEQUENCE 988 AA; 111783 MW; 0969BDECIAB43C4E CRC64;

Query Match 41.0%; Score 41; DB 1; Length 988;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 LKSTFCKPP 16
 ||:|||||
 DB 817 LGNTSTFCGP 828

RESULT 15
 KPC1_COCHHE STANDARD; PRT; 1174 AA.
 AC 042632;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein kinase C-like (EC 2.7.1.1-).
 GN PKC1.
 OS Cochliobolus heterostrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochliobolus.
 ON NCBI_TaxID=5016;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2 / ATCC 48329;
 RA Oeser B.M., Yoder O.C.;
 RT Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; Y15839; CAA75801.1; -;
 DR HSP; P05132; IAPM.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE_BIND.
 DR InterPro; IPR000719; Euk_kinase.
 DR InterPro; IPR000961; Kinas_C.
 DR InterPro; IPR000861; REM_repeat.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00130; DAG_PE-bind; 2.
 DR Pfam; PF02185; HRL; 2.
 DR Pfam; PF00069; Kinas_C; 1.
 DR Pfam; PF00433; Kinas_C; 1.
 DR SMART; SM00109; C1; 2.

DR SMART; SM00239; C2; 1.
 DR SMART; SM00074; HRL; 2.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; zinc;
 Phorbol-ester binding; Repeat.
 FT DOMAIN 459 506
 FT DOMAIN 527 576 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 704 709 POLY-ALA.
 FT DOMAIN 785 792 POLY-PRO.
 FT NP_BIND 849 1108 PROTEIN KINASE.
 FT BINDING 855 863 ATP (BY SIMILARITY).
 FT ACT_SITE 878 878 ATP (BY SIMILARITY).
 SQ SEQUENCE 1174 AA; 130506 MW; 26A4ADD42849F37C CRC64;

Query Match 41.0%; Score 41; DB 1; Length 1174;
 Best Local Similarity 44.4%; Pred. No. 57;
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 1 CASEL--GKSTFCKPP 16
 | | : | : | : | : |
 DB 996 CKEMMGSTSTFCGP 1013

Search completed: June 17, 2002, 16:24:28
 Job time: 1557 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:37 ; Search time 172.85 Seconds
(without alignments)
17.014 Million cell updates/sec

Title: US-09-761-636A-8
Perfect score: 100
Sequence: 1 CASELGKSTNTFCRPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL.19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query length	DB ID	Description
1	77.0	326	11	Q91ZE4	Q91ZE4 rat1us norv
2	59.0	326	11	Q91ZE6	Q91ZE6 meriones un
3	59.0	415	11	Q91ZE3	Q91ZE3 rat1us norv
4	59.0	418	13	Q57352	Q57352 colurnix co
5	59.0	420	6	Q9X550	Q9X550 bos taurus
6	50.0	889	11	Q9CUD6	Q9CUD6 mus musculu
7	49.0	127	13	Q918V8	Q918V8 rana pipien
8	48.5	746	10	Q91FT8	Q91FT8 arabidopsis
9	48.0	330	5	Q18118	Q18118 cseonorhadi
10	46.0	115	5	Q9NM00	Q9NM00 leishmania
11	45.0	367	10	Q9C6R6	Q9C6R6 arabidopsis
12	45.0	371	10	Q9C6F7	Q9C6F7 arabidopsis
13	45.0	442	11	Q62288	Q62288 mus musculu
14	44.0	145	5	Q9VE57	Q9VE57 drosophila
15	44.0	471	10	Q9FHK1	Q9FHK1 arabidopsis
16	44.0	677	10	Q9FLD9	Q9FLD9 arabidopsis

17	44.0	1189	5	Q9Y111	Q9Y111 drosophila
18	44.0	1189	5	Q9W378	Q9W378 drosophila
19	44.0	1717	5	Q26566	Q26566 schistosoma
20	43.0	100	16	Q92KW0	Q92KW0 helicobacte
21	43.0	263	10	Q9FL78	Q9FL78 arabidopsis
22	43.0	693	3	Q08001	Q08001 saccharomyc
23	43.0	849	4	Q96SQ3	Q96SQ3 homo sapien
24	43.0	997	5	Q967Z1	Q967Z1 trypanosoma
25	43.0	175	10	Q9LKW0	Q9LKW0 lycopersico
26	42.0	288	5	Q9XV5	Q9XV5 toxocara ca
27	42.0	326	10	Q93680	Q93680 persea amer
28	42.0	1011	3	Q9P944	Q9P944 pneumocysti
29	42.0	1428	5	Q44341	Q44341 halloctis ru
30	41.5	158	5	Q95V66	Q95V66 penaeus van
31	41.5	158	5	Q95V66	Q95V66 homo sapien
32	41.0	81	4	Q60400	Q60400 homo sapien
33	41.0	212	10	Q9M7H9	Q9M7H9 arabis alpi
34	41.0	219	5	Q44927	Q44927 toxocara ca
35	41.0	223	11	Q62220	Q62220 mus musculu
36	41.0	226	5	Q76131	Q76131 toxocara ca
37	41.0	230	11	Q64507	Q64507 mus musculu
38	41.0	252	10	Q9F120	Q9F120 arabidopsis
39	41.0	295	10	Q9M7G6	Q9M7G6 arabis gunn
40	41.0	316	13	Q902B2	Q902B2 brachydanio
41	41.0	320	10	Q9M7G7	Q9M7G7 arabis glab
42	41.0	335	10	Q957J5	Q957J5 arabidopsis
43	41.0	335	10	Q95838	Q95838 arabidopsis
44	41.0	335	10	Q95XJ4	Q95XJ4 arabidopsis
45	41.0	335	10	Q95XJ3	Q95XJ3 arabidopsis

ALIGNMENTS

RESULT	ID	Q91ZE4	PRELIMINARY:	PRT:	326 AA.
AC	Q91ZE4	Q91ZE4	Q91ZE4		
DT	01-DEC-2001	(TREMBlrel. 19, Created)			
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)			
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)			
DE	VEGF-D.				
OC	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SPRAGUE-DAWLEY;				
RA	Kirklin V., Mazitschek R., Krishnan J., Steffen A., Waltenberger J.,				
RT	Pepper M.S., Gianni A., Sleeman J.P.;				
RT	"Characterization of indolinones which specifically inhibit VEGF-C-and				
RL	VEGF-D-induced activation of VEGFR-3 but not VEGFR-2."				
RL	Eur. J. Biochem. 0:0-0(2001).				
DR	EMBL: AY032728; AAK96008.1;				
DR	SEQUENCE 326 AA; 37106 MW; D7CAEBA6C9FABD7D CRC64;				

Query Match	77.0%;	Score 77;	DB 11;	Length 326;
Best Local Similarity	87.5%;	Pred. No. 2.6e-05;		
Matches 14;	Conservative	1;	Mismatches 1;	Indels 0;
Gaps	0;			
OY	2	ASELGKSTNTFCRPPC 17		
DB	126	ASELGKSTNTFCRPPC 141		
RESULT	2			
ID	Q91ZE6	PRELIMINARY:	PRT:	326 AA.
AC	Q91ZE6			
DT	01-DEC-2001	(TREMBlrel. 19, Created)		
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)		

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C (FRAGMENT).
 GN VEGF.
 OS Meriones unguiculatus (Mongolian jird).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 RX NCBI_TaxID=10047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Panlatis B.J., Fuhrman J.A.;
 RT "Brugia malayi stimulates VEGF-C, a growth factor specific for
 lymphatic endothelium, by an indirect mechanism."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF432867; AAL28127.1; -
 FT NON-TER
 SQ SEQUENCE 326 AA; 36826 MW; D0B2772C7836914 CRC64;

Query Match
 Best Local Similarity 59.0%; Score 59; DB 11; Length 326;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTECKPPC 17
 DB 50 EFGAATNTEFKPPC 63

RESULT 3
 Q912E3 PRELIMINARY; PRT; 415 AA.
 AC Q912E3;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VEGF-C.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Kirkin V., Mazitschek R., Krishnan J., Steffen A., Waltenberger J.,
 RA Pepper M.S., Giannis A., Sleeman J.P.;
 RT "Characterization of endothelins which specifically inhibit VEGF-C and
 VEGF-D-induced activation of VEGFR-3 but not VEGFR-2."
 RL Eur. J. Biochem. 0:0-0(2001).
 DR EMBL: AY032729; AAK96009.1; -
 SO SEQUENCE 415 AA; 46397 MW; 1EB677F5B260A525 CRC64;

Query Match
 Best Local Similarity 59.0%; Score 59; DB 11; Length 415;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTECKPPC 17
 DB 139 EFGAATNTEFKPPC 152

RESULT 4
 057352 PRELIMINARY; PRT; 418 AA.
 AC 057352;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
 GN VEGF-C.
 OS Coturnix coturnix (common quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 RX NCBI_TaxID=9091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98167900; PubMed=9435294;
 RA Eichmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
 RA Allitalo K., Le Douarin N.M.;
 RT "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
 of the differentiation of VEGFR2 expressing endothelial cell
 precursors."
 RL Development 125:743-752(1998).
 DR EMBL: Y15837; CAA75799.1; -
 DR HSSP: P15692; 1VP.
 DR InterPro: IPR004153; CXXC.
 DR InterPro: IPR002400; GF_cyskn.
 DR InterPro: IPR000476; Glyco_hormone.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF03128; CXXC; 6.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PRODOM: PD001629; PDGF; 1.
 DR SMART: SM00067; GHA; 1.
 DR SMART: SM00141; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS0278; PDGF_2; 1.
 KW Signal.
 FT SIGNAL
 FT CHAIN 111 418 POTENTIAL.
 SQ SEQUENCE 418 AA; 46839 MW; 099BFC79151BF2B CRC64;

Query Match
 Best Local Similarity 59.0%; Score 59; DB 13; Length 418;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTECKPPC 17
 DB 142 EFGAATNTEFKPPC 155

RESULT 5
 Q9XS50 PRELIMINARY; PRT; 420 AA.
 AC Q9XS50;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RA MEDLINE=20044783; PubMed=10575000;
 RA Yonekura H., Sakurai S., Liu X., Migita H., Wang H., Yamagishi S.,
 RA Nomura M., Abedin M.J., Unoki H., Yamamoto Y., Yamamoto H.;
 RT "Placenta growth factor and vascular endothelial growth factor B and C
 expression in microvascular endothelial cells and pericytes.
 Implication in autocrine and paracrine regulation of angiogenesis."
 RL J. Biol. Chem. 274:35172-35178(1999).
 DR EMBL: AB004275; BAA77687.1; -
 DR HSSP: P15692; 1VP.
 DR InterPro: IPR004153; CXXC.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF03128; CXXC; 5.
 DR Pfam: PF00341; PDGF; 1.
 DR PRODOM: PD001629; PDGF; 1.
 DR SMART: SM00141; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.

DR PROSITE: PS50278; PDGF_2; 1.
 KW Signal. 1 20 POTENTIAL.
 FT SIGNAL. 21 420 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 FT CHAIN 1 420
 SQ SEQUENCE 420 AA; 46681 MW; 58BBA64317A3C8E2D CRC64;

Query Match 59.0%; Score 59; DB 6; Length 420;
 Best Local Similarity 71.4%; Pred. No. 0.044;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTFFCKPPC 17
 1 : |||||
 Db 144 EFGATNTFFKPPC 157

RESULT 6
 ID 09CUD6 PRELIMINARY; PRT: 889 AA.
 AC 09CUD6;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 4933403G14RIK PROTEIN (FRAGMENT).
 GN 4933403G14RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 ON NCBI_TaxID=10090;
 RX [1]

SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RC MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner U., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima U., Mazarrelli J., Mommaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK016632; BAB30348.1;
 DR MGD: MGI:1921643; 4933403G14RIK.
 FT NON_TER 889 889
 SQ SEQUENCE 889 AA; 95947 MW; AFE42D3971F1F954 CRC64;

Query Match 50.0%; Score 50; DB 11; Length 889;
 Best Local Similarity 64.3%; Pred. No. 3.3;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 SELGKSTNTFFCKPP 16
 1 : |||||
 Db 259 TELDEFNTFFCKPP 272

RESULT 7
 O918V8 PRELIMINARY; PRT: 127 AA.
 AC 0918V8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ONCOMASE VARIANT RAP1L PRECURSOR.
 OS Rana pipiens (Northern leopard frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
 ON NCBI_TaxID=8404;
 RN [1]

SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RC MEDLINE=20330357; PubMed=10871370;
 RA Chen S.-L., Le S.-Y., Newton D.L., Maizel J.V. Jr., Rybak S.M.;
 RT "A gender-specific mRNA encoding a cytotoxic ribonuclease contains a
 RT 3' UTR of unusual length and structure."
 RL Nucleic Acids Res. 28:2375-2382(2000).
 DR EMBL: AF165133; AAF6935.1;
 DR HSP: P22069; IONC.
 DR InterPro: IPR001427; RNaseA.
 DR Pfam: PF00074; RNaseA: 1.
 DR Prodom: PD000535; RNaseA: 1.
 DR SMART: SM00092; RNase_Pc: 1.
 DR PROSITE: PS00127; RNASE_PANCREATIC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL. 1 23 POTENTIAL.
 SQ SEQUENCE 127 AA; 14491 MW; B8511DC5407AB69B CRC64;

Query Match 49.0%; Score 49; DB 13; Length 127;
 Best Local Similarity 69.2%; Pred. No. 0.74;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFFC 13
 1 : |||||
 Db 98 CKYKLKSTNTFFC 110

RESULT 8
 ID 09LUT8 PRELIMINARY; PRT: 746 AA.
 AC 09LUT8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE GENOMIC DNA, CHROMOSOME 5, BAC CLONE: F1P19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RC Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB025603; BAA97466.1;
 DR MGD: MGI:1921643; 4933403G14RIK.
 FT NON_TER 746 746
 SQ SEQUENCE 746 AA; 82656 MW; A0EACBA409B5953A CRC64;

Query Match 48.5%; Score 48.5; DB 10; Length 746;
 Best Local Similarity 35.7%; Pred. No. 5.1;
 Matches 10; Conservative 3; Mismatches 4; Indels 11; Gaps 1;

OY 1 CASELGKSTNTFFC-----CKPPC 17
 1 : |||||
 Db 583 CSEVMEKSTNTFFSADFSTLADCKDPC 610

RESULT 9
 O18118 PRELIMINARY; PRT: 330 AA.
 ID O18118
 AC O18118;

```

DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN T23F1.6 PROTEIN.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: 281129; CAB03405.1;
DR InterPro; IPR003341; DUF139.
DR Pfam; PF02363; DUF139; 7.
SQ SEQUENCE 330 AA; 36605 MW; F043B1A90D3A8FE9 CRC64;

Query Match
Best Local Similarity 48.0%; Score 48; DB 5; Length 330;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 CASELGRSTNTECKPPC 17
Db 80 CSOCCSNTNTOCCPTC 96

RESULT 10
Q9NM00 PRELIMINARY; PRT; 115 AA.
ID Q9NM00;
AC Q9NM00;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE HYPOTHETICAL 12.7 KDA PROTEIN (FRAGMENT).
GN LM26.292.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=FRIDLIN.
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL160493; CAB97910.1;
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 115 AA; 12657 MW; 7E73C84D80B2B733 CRC64;

Query Match
Best Local Similarity 46.0%; Score 46; DB 5; Length 115;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 CASELGRSTNTECKPPC 17
Db 46 CAATAKSTISHCATPC 62

RESULT 11
Q9C6R6 PRELIMINARY; PRT; 367 AA.
ID Q9C6R6;
AC Q9C6R6;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

```

```

DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE HYPOTHETICAL 42.4 KDA PROTEIN (FRAGMENT).
GN T18124.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Elgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Millscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tatum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotska V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL: AC079131; AAG50757.1;
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PSS0181; FBOX; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 367 AA; 42444 MW; 751C6164205E7B4A CRC64;

Query Match
Best Local Similarity 45.0%; Score 45; DB 10; Length 367;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 ASELGKSTNTECKPPC 17
Db 220 SKIEKSRKTFVLP 235

RESULT 12
Q9C6F7 PRELIMINARY; PRT; 371 AA.
ID Q9C6F7;
AC Q9C6F7;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE HYPOTHETICAL 42.9 KDA PROTEIN.
GN T15M6.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Elgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,

```

RA Gil J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hopfer S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pail G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shin P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana,"
 RL Nature 408:816-820(2000).
 DR EMBL: AC079604; AAG50706.1;
 DR InterPro: IPR001810; F-box.
 DR Pfam: PF00646; F-box; 1.
 DR SMART: SM00256; FBOX; 1.
 DR PROSITE: PSS0181; FBOX; 1.
 DR Hypothetical protein.
 KW SEQUENCE 371 AA; 42890 MW; FCC77E57A1990457 CRC64;

Query Match 45.0%; Score 45; DB 10; Length 371;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ASELKSTNTECKRPPC 17
 Db 224 SKEIKSRKTECVLP 239

RESULT 13
 ID 062288 PRELIMINARY; PRT; 442 AA.
 AC 062288;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE SPOCK PROTEIN PRECURSOR (TESTICAN).
 GN SPOCK OR SPOCK1 OR TICN1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
 CC NCB1_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-BRAIN;
 RX MEDLINE-966224019; PubMed-8626787;
 RA Bonnet F., Perin J.P., Charbonnier F., Camuzat A., Roussel G.,
 RA Nussbaum J.L., Alliel P.M.;
 RT "Structure and cellular distribution of mouse brain testican.
 RT Association with the postsynaptic area of hippocampus pyramidal
 RT cells,"
 RL J. Biol. Chem. 271:4373-4380(1996).
 CC -1- FUNCTION: MAY PLAY A ROLE IN CELL-CELL AND CELL-MATRIX
 CC INTERACTIONS. MAY CONTRIBUTE TO VARIOUS NEURONAL MECHANISMS IN THE
 CC CENTRAL NERVOUS SYSTEM.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE POSTSYNAPTIC
 CC AREA OF PYRAMIDAL NEURONS.
 CC -1- PTM: CONTAINS CHONDROITIN SULFATE AND HEPARAN SULFATE O-LINKED
 CC OLIGOSACCHARIDES.
 CC -1- SIMILARITY: CONTAINS 4 OSTEOONECTIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.
 DR EMBL: X92864; CAA63448.1;
 DR MGD: MGI:105371; Spock1.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002350; kazal.
 DR InterPro: IPR000716; Thyroglobulin_1.
 DR Pfam: PF000050; kazal; 1.
 DR Pfam: PF00086; thyroglobulin_1; 1.
 DR SMART: SM00280; KAZAL; 1.

DR SMART: SM00211; TY; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS00484; THYROGLOBULIN_1; 1.
 KW Proteoglycan; Heparan sulfate; EGF-like domain; Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 442
 FT DOMAIN 18 50
 FT DOMAIN 64 111
 FT DOMAIN 142 169
 FT DOMAIN 252 308
 FT DOMAIN 313 379
 FT DOMAIN 389 429
 FT DOMAIN 386 386
 FT CARBOHYD 391 391
 FT SEQUENCE 442 AA; 49541 MW; 818C30313F8AC0F6 CRC64;

Query Match 45.0%; Score 45; DB 11; Length 442;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ASELKSTNTECKRPPC 17
 Db 168 ACSTGKSLNSLDCG 183

RESULT 14
 ID 09VE57 PRELIMINARY; PRT; 145 AA.
 AC 09VE57;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE CG7695 PROTEIN.
 GN CG7695.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CC NCB1_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blasej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Flannock C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003723; AAF55572.1; -;
 DR FlyBase: FBgn0038631; CG7695.
 DR InterPro: IPR002350; kazal.
 DR Pfam: PF00050; kazal.1.
 DR SMART: SM00280; KAZAL.1.
 SQ SEQUENCE 145 AA; 16581 MW; 4A52F2795CC67288 CRC64;

Query Match 44.0%; Score 44; DB 5; Length 145;
 Best Local Similarity 53.3%; Pred. No. 6.2;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 SELKSTNTCKPC 17
 |||::| ||| |
 Db 77 SELIETREICKPC 91

RESULT 15
 O9FHK1 PRELIMINARY; PRT; 471 AA.
 AC O9FHK1;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 3, P1 CLONE:MAB16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 OX 11
 RN SEQUENCE FROM N.A.
 RP STRAIN-COLUMBIA;
 RC MEDLINE=20181125; PubMed=10718197;
 RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RA "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones";
 RL DNA Res. 7:31-63(2000).
 DR EMBL: AB018112; BAB10975.1; -;
 DR EMBL: AB026661; BAB10975.1; JOINED.
 DR InterPro: IPR001810; F-box.
 DR Pfam: PF00646; F-box.1.
 DR SMART: SM00256; FBOX.1.
 DR PROSITE: PSS0181; FBOX.1.
 SQ SEQUENCE 471 AA; 55212 MW; 1E5D997B7942E26C CRC64;

Query Match 44.0%; Score 44; DB 10; Length 471;
 Best Local Similarity 58.3%; Pred. No. 19;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 ELCKSTNTCKPC 15
 |||::| ||| |
 Db 365 DLGKSTNTCKPC 376

Search completed: June 17, 2002, 16:23:38
 Job time: 1612 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:12 ; Search time 227.35 Seconds
(without alignments)
8.305 Million cell updates/sec

Title:	US-09-761-636A-8
Perfect score:	100
Sequence:	1 CASELGKSTNTECKPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

- 1: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-genseq/genseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100	100.0	17	22	AAU04527	VEGF based bicyclic
2	97	97.0	17	22	AAU04546	VEGF based bicyclic
3	97	97.0	17	22	AAU04550	VEGF based bicyclic
4	93	93.0	17	22	AAU04547	VEGF based bicyclic
5	93	93.0	17	22	AAU04549	VEGF based bicyclic
6	93	93.0	17	22	AAU04551	VEGF based bicyclic
7	89	89.0	17	22	AAU04548	VEGF based bicyclic
8	80	80.0	73	22	AAU04522	Human VEGF-D amino
9	80	80.0	96	22	AAU04520	Human VEGF-D amino
10	80	80.0	109	20	AAU23869	Human vascular endothelial
11	80	80.0	109	21	AAU11931	Human truncated VEGF

12	80	80	0	325	19	AAW53240	Homo sapiens vascu
13	80	80	0	325	22	AAV97572	Human VEGF-D prote
14	80	80	0	354	19	AAW49036	Human vegf2 growt
15	80	80	0	354	19	AAW53291	Homo sapiens vascu
16	80	80	0	354	19	AAW44233	Human vascular end
17	80	80	0	354	21	AAW10649	Human VEGD protein
18	80	80	0	354	21	AAW82904	Human VEGF-D prote
19	80	80	0	354	21	AAW70750	Human vegf-pro-vasc
20	80	80	0	354	21	AAW70983	Human vegf-pro-vasc
21	80	80	0	354	22	AAW08441	Polypeptide for hu
22	80	80	0	354	22	AAW08655	Human vascular end
23	80	80	0	354	22	AAV97563	Human VEGF-D1 prot
24	80	80	0	354	22	AAW37656	Human VEGF-D. Hom
25	80	80	0	620	18	AAW14994	Human c-Fos induce
26	77	77	0	178	20	AAW8287	Human growth facto
27	77	77	0	321	19	AAW53243	Homo sapiens vascu
28	77	77	0	326	19	AAW44296	Rat vascular endot
29	77	77	0	337	20	AAW08286	Human growth facto
30	77	77	0	358	18	AAW14992	Murine c-Fos induc
31	77	77	0	358	19	AAW53242	Mus musculus vascu
32	77	77	0	358	19	AAW44295	Mus vascular end
33	72	72	0	13	22	AAU04534	VEGF based monocyc
34	69	69	0	13	22	AAU04534	VEGF based monocyc
35	65	65	0	13	22	AAU04535	VEGF based monocyc
36	65	65	0	13	22	AAU04537	VEGF based monocyc
37	61	61	0	13	22	AAU04536	VEGF based monocyc
38	59	59	0	415	18	AAW09933	Mouse Flk4 recepto
39	59	59	0	415	19	AAW75742	Mouse vascular end
40	59	59	0	418	18	AAW00934	Quail Flk4 recepto
41	59	59	0	418	19	AAW75743	Quail vascular end
42	58	58	0	21	22	AAU08479	Peptide #1 inhibit
43	58	58	0	21	22	AAU08480	Peptide #2 inhibit
44	58	58	0	21	22	AAU08485	VEGFR-3 binding ep
45	58	58	0	54	22	AAU08422	Polypeptide encode

ALIGNMENTS

RESULT	1
AAU04527	standard; Protein; 17 AA.
AC	AAU04527;
AD	26-SEP-2001 (first entry)
AE	VEGF based bicyclic dimeric peptide #1.
AF	Human: VEGF: vascular endothelial growth factor; angiogenesis;
AG	neovascularisation; lymphangiogenesis; psoriasis; tumour;
AH	KW diabetic induced neovascular sequelae; rheumatoid arthritis;
AI	diabetic retinopathy; chronic inflammation; cyclic.
AK	Synthetic.
AL	Key
AM	Disulfide-bond
AN	1..13
AO	/note= "this bond cyclises the peptide"
AP	Disulfide-bond
AQ	17
AR	/note= "A disulfide bond forms between residue
AS	17 and residue 17 of an identical peptide to form
AT	a dimeric peptide, or to residue 1 of the sequence
AW	appearing as AAU04528, also forming a dimeric peptide
AX	WO200152875-A1.
AY	26-JUL-2001.
AZ	18-JAN-2001; 2001WO-US01533.
BA	18-JAN-2000; 2000US-0176283.
BB	16-MAY-2000; 2000US-0204590.
BC	
BD	
BE	
BF	
BG	
BH	
BI	
BJ	
BK	
BL	
BM	
BN	
BO	
BP	
BQ	
BR	
BS	
BT	
BU	
BV	
BW	
BX	
BY	
BZ	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	

XX (LUDW-) LUDWIG INST CANCER RES.
 PA Achen MG, Hughes RA, Stacker S, Cendron A;
 PI WPI: 2001-442248/47.
 XX
 DR WPI: 2001-442248/47.
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclizing a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Claim 59; Page 32; 102pp: English.
 XX
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the exposed loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC a beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclizing the peptide by oxidizing the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 XX
 SO Sequence 17 AA:
 QY Query Match 100.0%; Score 100; DB 22; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 CASLGGKSTNTFCRPPC 17
 1 caselgkstntfcckpc 17
 RESULT 2
 AA004546
 ID AA004546 standard; Peptide: 17 AA.
 AC AA004546;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based bicyclic dimeric peptide #3.
 XX
 KM Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation; cyclic.
 XX
 OS Synthetic.
 XX
 Key Location/Qualifiers
 FH Disulfide-bond 1..13
 FT

FT Disulfide-bond 17 /note="This bond cyclises the peptide"
 FT 17 /note="A disulfide bond forms between residue
 FT 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide"
 XX
 PN WO200152875-A1.
 XX
 PD 26-JUL-2001.
 XX
 XX 18-JAN-2001; 2001WO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 PI WPI: 2001-442248/47.
 XX
 PS Example 26; Page 49; 102pp: English.
 XX
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the exposed loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC a beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclizing the peptide by oxidizing the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 XX
 SO Sequence 17 AA:
 QY Query Match 97.0%; Score 97; DB 22; Length 17;
 Best Local Similarity 94.1%; Pred. No. 8e-08;
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 DB 1 CASLGGKSTNTFCRPPC 17
 1 caselgkstntfcckpc 17
 RESULT 3
 AA004550
 ID AA004550 standard; Peptide: 17 AA.
 AC AA004550;
 XX

XX 26-SEP-2001 (first entry)
 XX VEGF based bicyclic dimeric peptide #7.
 DE Human: VEGF; vascular endothelial growth factor; angiogenesis;
 XX neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..13
 FT /note= "This bond cyclises the peptide"
 FT Disulfide-bond 17
 FT /note= "A disulfide bond forms between residue
 FT 17 and residue 1 of the sequence appearing as AAU04528,
 FT forming a dimeric peptide"
 FT
 PN W0200152875-A1.
 XX 26-JUL-2001.
 PD 18-JAN-2001; 2001WO-US01533.
 XX 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA Achen MG, Hughes RA, Stacker S, Cendron A;
 PI WPI: 2001-442248/47.
 DR Novel monomeric monocyclic peptide, used to interfere with
 XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Example 26; Page 49; 102pp; English.
 XX
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides), dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 XX Sequence 17 AA;

Query Match 97.0%; Score 97; DB 22; Length 17;
 Best Local Similarity 94.1%; Pred. No. 8e-08;
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CASELGKSTWTFCKKPC 17
 DB 1 caseLgkstntcyckpc 17
 RESULT 4
 AAU04547
 ID AAU04547 standard; Peptide: 17 AA.
 AC AAU04547;
 XX 26-SEP-2001 (first entry)
 DE VEGF based bicyclic dimeric peptide #4.
 XX Human: VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..13
 FT /note= "This bond cyclises the peptide"
 FT Disulfide-bond 17
 FT /note= "A disulfide bond forms between residue
 FT 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide"
 FT
 PN W0200152875-A1.
 XX 26-JUL-2001.
 PD 18-JAN-2001; 2001WO-US01533.
 XX 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA Achen MG, Hughes RA, Stacker S, Cendron A;
 PI WPI: 2001-442248/47.
 DR Novel monomeric monocyclic peptide, used to interfere with
 XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Example 26; Page 49; 102pp; English.
 XX
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGFD (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides), dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular

CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
SQ Sequence 17 AA;

Query Match 93.0%; Score 93; DB 22; Length 17;
Best Local Similarity 88.2%; Pred. No. 3.2e-07;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFCCKPCC 17
DB 1 caseLgkstntfcckppc 17
|||||:|||||

RESULT 5
AAU04549 ID AAU04549 standard; Peptide: 17 AA.
XX
AC AAU04549;

DT 26-SEP-2001 (first entry)

DE VEGF based bicyclic dimeric peptide #6.

XX
XX Human: VEGF: vascular endothelial growth factor; angiogenesis;
KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
XX diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..13 /note- "this bond cyclises the peptide"

FT Disulfide-bond 17 /note- "A disulfide bond forms between residue
FT 17 and residue 17 of an identical peptide to form
XX a dimeric peptide"

PN WO200152875-A1.

XX 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01533.

PR 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Achen MG, Hughes RA, Stacker S, Cendron A;

DR WPI: 2001-442248/47.

XX
XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

PS Example 26; Page 49; 102pp; English.

XX The sequence represents a dimeric bicyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGFD (vascular endothelial growth factor). The invention relates to a

CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, heart or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
SQ Sequence 17 AA;

Query Match 93.0%; Score 93; DB 22; Length 17;
Best Local Similarity 88.2%; Pred. No. 3.2e-07;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CASELGKSTNTFCCKPCC 17
DB 1 caseLgkstntfcckppc 17
|||||:|||||

RESULT 6
AAU04551 ID AAU04551 standard; Peptide: 17 AA.
XX
AC AAU04551;

DT 26-SEP-2001 (first entry)

DE VEGF based bicyclic dimeric peptide #8.

XX
XX Human: VEGF: vascular endothelial growth factor; angiogenesis;
KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
XX diabetic retinopathy; chronic inflammation; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..13 /note- "this bond cyclises the peptide"

FT Disulfide-bond 17 /note- "A disulfide bond forms between residue
FT 17 and residue 1 of the sequence appearing as AAU04528,
XX forming a dimeric peptide"

PN WO200152875-A1.

XX 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01533.

PR 18-JAN-2000; 2000US-0176293.

PR 16-MAY-2000; 2000US-0204590.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Achen MG, Hughes RA, Stacker S, Cendron A;
 XX
 XX WPI: 2001-442248/47.
 DR
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 XX
 PS Example 26; Page 49; 102pp; English.
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 XX
 SQ Sequence 17 AA:
 Query Match 93.0%; Score 93; DB 22; Length 17;
 Best Local Similarity 88.2%; Pred. No. 3.2e-07;
 Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CASELGKSTNTPCKPPC 17
 |||||:||||:|||||
 Db 1 caselgrstnsfckppc 17
 RESULT 7
 AAU04548
 ID AAU04548 standard; Peptide: 17 AA.
 XX
 XX AAU04548;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX VEGF based bicyclic dimeric peptide #5.
 DE
 XX Human; VEGF: vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.
 XX
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..13
 FT /note- "This bond cyclises the peptide"
 FT Disulfide-bond 17
 FT /note- "A disulfide bond forms between residue

FF 17 and residue 17 of an identical peptide to form
 FT a dimeric peptide"
 XX
 XX WO200152875-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 18-JAN-2001; 2001WO-US01533.
 XX
 XX 18-JAN-2000; 2000US-0176293.
 XX
 XX 16-MAY-2000; 2000US-0204590.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Achen MG, Hughes RA, Stacker S, Cendron A;
 PI
 XX WPI: 2001-442248/47.
 DR
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 XX
 PS Example 26; Page 49; 102pp; English.
 CC The sequence represents a dimeric bicyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 XX
 SQ Sequence 17 AA:
 Query Match 89.0%; Score 89; DB 22; Length 17;
 Best Local Similarity 82.4%; Pred. No. 1.3e-06;
 Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CASELGKSTNTPCKPPC 17
 |||||:|||||:|||||
 Db 1 casdvgrstntwckppc 17
 RESULT 8
 AAU04522
 ID AAU04522 standard; protein: 73 AA.
 XX
 XX AAU04522;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX

DE Human VEGF-D amino acids Val101-Thr 173.
 XX
 KM Human: VEGF-D; vascular endothelial growth factor; angiogenesis;
 KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation.
 XX
 OS Homo sapiens.
 PN WO200152875-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 XX
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 XX
 DR WPI: 2001-442248/47.
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Example 1: Page 90-91; 102pp; English.
 XX
 CC The sequence represents Human VEGF-D (vascular endothelial growth factor)
 CC amino acids Val101-Thr 173, used together with the C-terminal 23
 CC residues of VEGF to make a hybrid theoretical molecule for 3
 CC dimensional modelling. The sequence is used in a method of producing
 CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
 CC distances on opposite antiparallel strands of a peptide loop fragment
 CC from an exposed loop of a growth factor protein and cyclising the peptide
 CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
 CC bicyclic peptides (comprising 2 linked monocyclic peptides), and a cyclic
 CC peptide with at least one amino acid deleted prior to cyclisation are
 CC used to interfere with angiogenesis, neovascularisation or
 CC lymphangiogenesis in a mammal with a condition characterised by
 CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
 CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
 CC malignant or benign tumour, post-recovery cerebrovascular accident,
 CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
 CC neovascularisation of the liver, excessive hormone-related angiogenic
 CC dysfunction, diabetes induced neovascular sequelae, hypertension-induced
 CC neovascular sequelae, or chronic liver infection. The peptides are also
 CC used to modulate vascular permeability in a mammal (the mammal has a
 CC condition characterised by fluid accumulation in peripheral limbs or in
 CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
 CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
 CC peptides are used to interfere with at least one biological activity
 CC induced by VEGF, VEGF-C or -D and are also used in combination with an
 CC anti-inflammatory agent, to treat a chronic inflammation, especially a
 CC rheumatoid arthritis, psoriasis and diabetic retinopathy.
 XX
 SQ Sequence 73 AA;

Query Match 80.0%; Score 80; DB 22; Length 73;
 Best Local Similarity 93.8%; Pred. No. 0.00011;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ASELGKSTNTFFCKPPC 17
 |||||
 Db 21 aselgkstntffkppc 36

ID AAU04520 standard; protein; 96 AA.
 XX
 AC AAU04520;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Human VEGF-D amino acids Val101-PRO186.
 XX
 KM Human: VEGF-D; vascular endothelial growth factor; angiogenesis;
 KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation.
 XX
 OS Homo sapiens.
 PN WO200152875-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 XX
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 XX
 DR WPI: 2001-442248/47.
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Example 1: Page 89; 102pp; English.
 XX
 CC The sequence represents Human VEGF-D (vascular endothelial growth factor)
 CC amino acids Val101-PRO186. The sequence is used in a method of producing
 CC a monomeric monocyclic peptide by a measuring beta-beta carbon separation
 CC distances on opposite antiparallel strands of a peptide loop fragment
 CC from an exposed loop of a growth factor protein and cyclising the peptide
 CC by oxidising the cysteine residues. The monocyclic peptides, dimeric
 CC bicyclic peptides (comprising 2 linked monocyclic peptides), and a cyclic
 CC peptide with at least one amino acid deleted prior to cyclisation are
 CC used to interfere with angiogenesis, neovascularisation or
 CC lymphangiogenesis in a mammal with a condition characterised by
 CC angiogenesis, neovascularisation or lymphangiogenesis. The condition is
 CC diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
 CC malignant or benign tumour, post-recovery cerebrovascular accident,
 CC post-angioplasty restenosis, head, heat or cold trauma, substance-induced
 CC neovascularisation of the liver, excessive hormone-related angiogenic
 CC dysfunction, diabetes induced neovascular sequelae, hypertension-induced
 CC neovascular sequelae, or chronic liver infection. The peptides are also
 CC used to modulate vascular permeability in a mammal (the mammal has a
 CC condition characterised by fluid accumulation in peripheral limbs or in
 CC lungs, peritoneal cavity, pleura, or brain. The peptides are used to
 CC image blood vessels and lymphatic vasculature. The monomeric and bicyclic
 CC peptides are used to interfere with at least one biological activity
 CC induced by VEGF, VEGF-C or -D and are also used in combination with an
 CC anti-inflammatory agent, to treat a chronic inflammation, especially a
 CC rheumatoid arthritis, psoriasis and diabetic retinopathy.
 XX
 SQ Sequence 96 AA;

Query Match 80.0%; Score 80; DB 22; Length 96;
 Best Local Similarity 93.8%; Pred. No. 0.00015;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ASELGKSTNTFFCKPPC 17
 |||||
 Db 21 aselgkstntffkppc 36

RESULT 10
 AAY23889
 ID AAY23889 standard; Protein: 109 AA.
 XX
 AC AAY23889;
 XX
 DT 21-SEP-1999 (first entry)
 XX
 DE Human vascular endothelial growth factor (VEGF)-D.
 XX
 KW Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
 KW tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
 KW wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
 XX
 OS Homo sapiens.
 XX
 PN MO9933485-A1.
 XX
 PD 08-JUL-1999.
 XX
 PF 23-DEC-1998; 98WO-US27373.
 XX
 PR 29-MAY-1998; 98US-0087392.
 PR 24-DEC-1997; 97AU-0001131.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Alitalo K, Stacker SA;
 XX
 DR WPI: 1999-405368/34.
 XX
 PT A human cell line stably expressing vascular endothelial growth
 factor D, useful for treating melanomas or tumours expressing VEGF-D
 XX
 PS Claim 6; Page 72; 79pp; English.
 XX
 CC The present sequence represents human vascular endothelial growth
 factor (VEGF)-D. The specification describes a human cell line
 CC which stably expresses VEGF-D, or fragments/analogs having VEGF-D
 CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
 CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
 CC antibodies), are useful for the treatment or alleviation of malignant
 CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
 CC stimulating amounts of VEGF-D can be administered to enhance the
 CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
 CC acceptance and/or healing of skin grafts or to stimulate the healing
 CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
 CC stimulating amounts of VEGF-D can be used to treat lymphedema.
 CC Endothelial proliferation stimulating amounts of VEGF-D are used to
 CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
 CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
 CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
 CC can be used to stimulate at least one VEGF-D bioactivity chosen from
 CC endothelial cell proliferation, migration, survival and differentiation
 CC and lymphangiogenesis without inducing vascular permeability.
 CC
 XX
 SQ Sequence 109 AA:
 XX
 QY
 DB 29 aselgkstntffkppc 44

Query Match 80.0%; Score 80; DB 20; Length 109;
 Best Local Similarity 93.8%; Pred. No. 0.00016;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
 AAB11931
 ID AAB11931 standard; Protein: 109 AA.
 XX

AC AAB11931;
 XX
 DT 20-NOV-2000 (first entry)
 XX
 DE Human truncated VEGF-D.
 XX
 KW Truncated VEGF-D; vascular endothelial growth factor; human;
 KW monoclonal antibody; VEGF receptor; VEGFR-2; VEGFR-3;
 KW vascular permeability disorder; endothelial cell proliferative disorder;
 KW angiogenic disorder; lymphangiogenic disorder;
 KW neovascularisation disorder; endothelial cell differentiation disorder;
 KW cancer; diabetic retinopathy; psoriasis; arthropathy; pulmonary oedema;
 KW detection; diagnosis; imaging; lymphatic vasculature.
 XX
 OS Homo sapiens.
 XX
 PN WO200037025-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 21-DEC-1999; 99WO-US31332.
 XX
 PR 21-DEC-1998; 98US-0113254.
 PR 17-MAY-1999; 99US-0134556.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Stacker SA;
 XX
 DR WPI: 2000-442498/38.
 XX
 PT Novel compositions comprising antibodies reactive to vascular
 endothelial growth factor-D, useful for treating, e.g. angiogenesis,
 PT lymphangiogenesis and neovascularization disorders -
 XX
 PS Claim 1; Fig 1; 44pp; English.
 XX
 CC This sequence represents a 109 amino acid truncated human VEGF-D
 CC (vascular endothelial growth factor D), lacking both the N- and
 CC C-terminal regions. The invention relates to a monoclonal antibody, or
 CC fragments thereof, which is specifically reactive with the truncated
 CC human VEGF-D, and methods of preparing the antibody. The antibody of the
 CC invention interferes with the binding of VEGF-D to the VEGF receptors
 CC VEGFR-2 and VEGFR-3, but does not interfere with the binding of VEGF to
 CC these receptors and additionally is not reactive with VEGF-C. The
 CC antibody may be used to treat disorders associated with vascular
 CC permeability, endothelial cell proliferation, angiogenesis,
 CC lymphangiogenesis, neovascularisation and endothelial cell
 CC arthropathies. The antibody may also be used to treat fluid accumulation
 CC in the heart and/or lung via modulation of vascular permeability. It may
 CC additionally be used to detect VEGF-D and may be used to image lymphatic
 CC vasculature in tissue.
 CC
 XX
 SQ Sequence 109 AA:
 XX
 QY
 DB 29 aselgkstntffkppc 44

Query Match 80.0%; Score 80; DB 21; Length 109;
 Best Local Similarity 93.8%; Pred. No. 0.00016;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
 AAW53240
 ID AAW53240 standard; Protein: 325 AA.
 XX
 AC AAW53240;
 XX
 DT 03-AUG-1998 (first entry)
 XX

XX Sequence 325 AA;
SQ
Query Match 80.0%; Score 80; DB 22; Length 325;
Best Local Similarity 93.8%; Pred. No. 0.00047;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 ASELGKSTNTFCCKPCC 17
|||||
DB 92 aselgkstntffkppc 107

RESULT 14
AAM49036
ID AAM49036 standard; Protein: 354 AA.
XX
AC AAM49036;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human zvegf2 growth factor.
XX
KW Human zvegf2 growth factor; mitogen; fibroblast; smooth muscle cell;
KW venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
KW angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
KW scleroderma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Signal peptide"
FT Peptide 24..108
FT /note= "Pro-region"
FT Binding-site 109..197
FT /note= "Receptor binding domain"
FT Region 206..256
FT /note= "Cysteine-rich domain"
FT Region 257..274
FT /note= "Balbiani ring motif"
FT Region 275..294
FT /note= "Balbiani ring motif"
FT Region 295..354
FT /note= "Cysteine-rich domain"
XX
PN MO9824811-A2.
XX
PD 11-JUN-1998.
XX
PF 20-NOV-1997; 97WO-US20888.
XX
PR 18-SEP-1997; 97US-0933455.
PR 06-DEC-1996; 96US-0759657.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC, Gilbert T, Hart CE, Nygaard S, Sheppard PO;
XX
DR WPI: 1998-333256/29.
DR N-PSDB; AAV32823.
XX
PT New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX
XX Claim 1; Pages 53-54; 77pp; English.
XX
CC The present sequence represents a human zvegf2 growth factor encoded
CC by the zvegf2 cDNA which was isolated from a human heart cDNA library.
CC zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
CC smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
CC revascularisation of tissue or the re-endothelialisation of vascular

CC tissue. zvegf2 is particularly claimed to be useful for the treatment
CC of full-thickness skin wounds, including venous stasis ulcers and
CC diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
CC additive in tissue adhesives for promoting revascularisation of the
CC healing tissue. Antagonists against zvegf2 can be used to block its
CC mitogenic, chemotactic and angiogenic effects. The antagonists may
CC therefore be useful for reducing growth of solid tumours by inhibiting
CC neovascularisation of the developing tumour or by directly blocking
CC tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
CC arthritis, and scleroderma.
XX
XX Sequence 354 AA;
SQ
Query Match 80.0%; Score 80; DB 19; Length 354;
Best Local Similarity 93.8%; Pred. No. 0.00051;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 ASELGKSTNTFCCKPCC 17
|||||
DB 121 aselgkstntffkppc 136

RESULT 15
AAM53241
ID AAM53241 standard; Protein: 354 AA.
XX
AC AAM53241;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Homo sapiens.
XX
PN MO9807832-A1.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14696.
XX
PR 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
XX (LUDWIG INST CANCER RES.
XX (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Achen MG, Altalo K, Stacker SA, Wilks AF;
XX
DR WPI: 1998-179057/16.
DR N-PSDB; AAV20807.
XX
PT New isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
XX Claim 16; Pages 60-61; 101pp; English.

CC The sequence is that of human lung vascular endothelial growth factor
 CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
 CC in wound healing, tissue or organ transplantation, or to establish
 CC collateral circulation in tissue infarction or arterial stenosis,
 CC such as coronary artery disease, and inhibition of angiogenesis in
 CC the treatment of cancer or of diabetic retinopathy. It can also be
 CC used in the treatment of lung disorders to improve blood circulation
 CC in the lung and/or gaseous exchange between the lungs and the blood
 CC stream or to improve blood circulation to the heart and O2 gas
 CC permeability in cases of cardiac insufficiency, to improve blood
 CC flow and gaseous exchange in chronic obstructive airway disease,
 CC or to treat malabsorptive syndromes in the intestinal tract.
 CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
 CC as an indicator of future metastatic risk. Antagonists can be used
 CC for treating e.g. conditions such as congestive heart failure,
 CC involving accumulations of fluid in the lung resulting from
 CC increases in vascular permeability. The products can also be used
 CC for detection and diagnosis.
 CC
 XX Sequence 354 AA:

Query Match 80.0%; Score 80; DB 19; Length 354;
 Best Local Similarity 93.8%; Pred. NO. 0.00051;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ASELGKSTNTFCCKPPC 17
 |||||
 Db 121 aselgkstntfckppc 136

Search completed: June 17, 2002, 16:02:12
 Job time: 421 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:36 ; Search time 75.98 Seconds
(without alignments)
5.465 Million cell updates/sec

Title: US-09-761-636A-8

Perfect score: 100

Sequence: 1 CASELGKSTNFCCKPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	80.0	325	4	US-08-915-795-3 Sequence 3, Appli
2	80	80.0	354	4	US-08-915-795-5 Sequence 5, Appli
3	77	77.0	321	4	US-08-915-795-9 Sequence 9, Appli
4	77	77.0	358	4	US-08-915-795-8 Sequence 8, Appli
5	59	59.0	415	4	US-08-795-430-11 Sequence 11, Appli
6	59	59.0	418	4	US-08-795-430-13 Sequence 13, Appli
7	58	58.0	350	2	US-08-999-811-4 Sequence 4, Appli
8	58	58.0	350	2	US-08-824-996-2 Sequence 2, Appli
9	58	58.0	350	3	US-09-042-105-4 Sequence 4, Appli
10	58	58.0	350	4	US-08-510-133A-33 Sequence 33, Appli
11	58	58.0	350	4	US-08-585-895-33 Sequence 33, Appli
12	58	58.0	419	2	US-08-999-811-2 Sequence 2, Appli
13	58	58.0	419	3	US-09-042-105-2 Sequence 2, Appli
14	58	58.0	419	3	US-09-042-105-18 Sequence 18, Appli
15	58	58.0	419	4	US-08-795-430-8 Sequence 8, Appli
16	58	58.0	419	4	US-08-510-133A-35 Sequence 35, Appli
17	58	58.0	419	5	PCT-US96-09001-2 Sequence 2, Appli
18	47	47.0	120	2	US-08-938-975-2 Sequence 2, Appli
19	46	46.0	260	4	US-09-162-021B-4 Sequence 4, Appli
20	45	45.0	20	4	US-09-230-222-28 Sequence 28, Appli
21	43	43.0	83	3	US-08-875-811-2 Sequence 2, Appli
22	43	43.0	104	1	US-08-283-971-1 Sequence 1, Appli
23	43	43.0	104	1	US-07-921-619-1 Sequence 1, Appli
24	43	43.0	104	1	US-08-467-955-1 Sequence 1, Appli
25	43	43.0	104	1	US-08-467-955-2 Sequence 2, Appli
26	43	43.0	104	3	US-08-891-848-13 Sequence 13, Appli
27	43	43.0	104	2	US-08-875-811-1 Sequence 1, Appli

28	43	43.0	104	4	US-09-394-268-1 Sequence 1, Appli
29	43	43.0	104	4	US-09-394-268-2 Sequence 2, Appli
30	43	43.0	105	4	US-08-875-811-24 Sequence 24, Appli
31	43	43.0	105	3	US-08-875-811-26 Sequence 26, Appli
32	43	43.0	105	3	US-08-875-811-39 Sequence 39, Appli
33	43	43.0	106	3	US-08-875-811-28 Sequence 28, Appli
34	43	43.0	107	3	US-08-875-811-20 Sequence 20, Appli
35	43	43.0	107	3	US-08-875-811-30 Sequence 30, Appli
36	43	43.0	112	3	US-08-875-811-22 Sequence 22, Appli
37	43	43.0	112	3	US-08-875-811-32 Sequence 32, Appli
38	43	43.0	129	3	US-08-875-811-63 Sequence 63, Appli
39	43	43.0	251	3	US-08-875-811-59 Sequence 59, Appli
40	43	43.0	254	3	US-08-875-811-61 Sequence 61, Appli
41	43	43.0	355	3	US-08-875-811-41 Sequence 41, Appli
42	43	43.0	355	3	US-08-875-811-49 Sequence 49, Appli
43	43	43.0	355	3	US-08-875-811-57 Sequence 57, Appli
44	43	43.0	355	3	US-08-875-811-64 Sequence 64, Appli
45	43	43.0	358	3	US-08-875-811-51 Sequence 51, Appli

ALIGNMENTS

RESULT 1
US-08-915-795-3
; Sequence 3, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Kari ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Breast
; US-08-915-795-3
Query Match 80.0%; Score 80; DB 4; Length 325;
Best Local Similarity 93.8%; Pred. No. 8.6e-05;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFFCKPPC 17
|||||
Db 92 ASELGKSTNTFFCKPPC 107

RESULT 2

US-08-915-795-5
; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, Mckeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
US-08-915-795-5

Query Match

Best Local Similarity 80.0%; Score 80; DB 4; Length 354;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFFCKPPC 17
|||||
Db 121 ASELGKSTNTFFCKPPC 136

RESULT 3

US-08-915-795-9
; Sequence 9, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO

; TITLE OF INVENTION: GROWTH FACTOR

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, Mckeown, Edwards & Lenahan P.L.L.C.

; STREET: 1200 G Street, NW, Suite 700

; CITY: Washington

; STATE: DC

; COUNTRY: United States of America

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/915,795

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 1064/42983

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; TELEX: N/A

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 321 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; TISSUE TYPE: Mouse Lung

US-08-915-795-9

Query Match

Best Local Similarity 77.0%; Score 77; DB 4; Length 321;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ASELGKSTNTFFCKPPC 17
|||||
Db 121 ASELGKSTNTFFCKPPC 136

RESULT 4

US-08-915-795-8
; Sequence 8, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, Mckeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Mouse Lung
TISSUE TYPE: Mouse Lung
US-08-915-795-8

Query Match 77.0% Score 77; DB 4; Length 358;
Best Local Similarity 87.5%; Pred. No. 0.00027;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ASELGKSTNTFFCKPPC 17
Db 126 ASELGKSTNTFFCKPPC 141

RESULT 5
US-08-795-430-11
Sequence 11, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aitalo, Kari
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/ET96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011

FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-11

Query Match 59.0% Score 59; DB 4; Length 415;
Best Local Similarity 71.4%; Pred. No. 0.17;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ELGKSTNTFFCKPPC 17
Db 139 EFGATNTFFCKPPC 152

RESULT 6
US-08-795-430-13
Sequence 13, Application US/08795430
Patent No. 6130071
GENERAL INFORMATION:
APPLICANT: Aitalo, Kari
TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/ET96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.

REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-430-13

Query Match 59.0%; Score 59; DB 4; Length 418;
Best Local Similarity 71.4%; Pred. No. 0.17;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELGKSTNTECKPPC 17
| : |||| ||||
Db 142 EFGVATNTEFKPPC 155

RESULT 7
US-08-999-811-4
Sequence 4, Application US/08999811
Patent No. 5932540

GENERAL INFORMATION:

APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.

APPLICANT: CAO, LIANG

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVENUE

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/999,811

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/465,968

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MARKOWICZ, KAREN R.

REGISTRATION NUMBER: 36,351

REFERENCE/DOCKET NUMBER: 1488,1000004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)371-2600

TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 350 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-999-811-4

Query Match 58.0%; Score 58; DB 2; Length 350;

Best Local Similarity 71.4%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELGKSTNTECKPPC 17
| : |||| ||||
Db 74 EFGVATNTEFKPPC 87

RESULT 8

US-08-824-996-2

Sequence 2, Application US/08824996B

Patent No. 5935820

GENERAL INFORMATION:

APPLICANT: Hu, Jing-Shan

APPLICANT: Rosen, Craig A.

APPLICANT: Cao, Liang

TITLE OF INVENTION: Polynucleotides Encoding Vascular Endothelial Growth

FILE OF INVENTION: Factor 2

FILE REFERENCE: P112D1

CURRENT APPLICATION NUMBER: US/08/824,996B

CURRENT FILING DATE: 1997-03-27

EARLIER APPLICATION NUMBER: 08/207,550

EARLIER FILING DATE: 1994-03-08

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 2

LENGTH: 350

TYPE: PRT

ORGANISM: Homo sapiens

US-08-824-996-2

Query Match 58.0%; Score 58; DB 2; Length 350;
Best Local Similarity 71.4%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELGKSTNTECKPPC 17
| : |||| ||||
Db 74 EFGVATNTEFKPPC 87

RESULT 9

US-09-042-105-4

Sequence 4, Application US/09042105

Patent No. 6040157

GENERAL INFORMATION:

APPLICANT: HU, JING-SHAN

APPLICANT: ROSEN, CRAIG A.

APPLICANT: CAO, LIANG

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVENUE

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,105

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/465,968

FILED DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2680
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-4

Query Match 58.0%; Score 58; DB 3; Length 350;
Best Local Similarity 71.4%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ELKSTNTEFKPPC 17
Db 74 EFGVATNTEFKPPC 87

RESULT 10
US-08-510-133A-33
Sequence 33, Application US/08510133A
Patent No. 6221839
GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510.133A
FILING DATE: 01-Aug-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/332863
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-08-510-133A-33

Query Match 58.0%; Score 58; DB 4; Length 350;
Best Local Similarity 71.4%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ELKSTNTEFKPPC 17
Db 74 EFGVATNTEFKPPC 87

RESULT 11
US-08-585-895-33
Sequence 33, Application US/08585895
Patent No. 6245530
GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
Joukov, Vladimir
TITLE OF INVENTION: Receptor Ligand
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,895
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33072
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-585-895-33

Query Match 58.0%; Score 58; DB 4; Length 350;
Best Local Similarity 71.4%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ELKSTNTEFKPPC 17
Db 74 EFGVATNTEFKPPC 87

RESULT 12
US-08-999-811-2
Sequence 2, Application US/08999811
Patent No. 5932540
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
APPLICANT: CAO, LIANG
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,811
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-811-2

Query Match 58.0%; Score 58; DB 2; Length 419;
Best Local Similarity 71.4%; Pred. No. 0.24;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTFCKPPC 17
| | : ||| | |||
DB 143 EFGVATNTFFKPPC 156

RESULT 13
US-09-042-105-2
Sequence 2, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550

FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: ERIC K. STEFFE
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-105-2

Query Match 58.0%; Score 58; DB 3; Length 419;
Best Local Similarity 71.4%; Pred. No. 0.24;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 ELKSTNTFCKPPC 17
| | : ||| | |||
DB 143 EFGVATNTFFKPPC 156

RESULT 14
US-09-042-105-18
Sequence 18, Application US/09042105
Patent No. 6040157
GENERAL INFORMATION:
APPLICANT: HU, JING-SHAN
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVENUE
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,105
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,550
FILING DATE: 8-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TO BE ASSIGNED
FILING DATE: 24-DEC-1997
CLASSIFICATION:

NAME: Gass, David A.

Search completed: June 17, 2002, 15:56:37
Job time: 86 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:22 ; Search time 95.43 Seconds
(without alignments)
9.062 Million cell updates/sec

Title: US-09-761-636a-9

Sequence: 1 CCNESLIC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	81.8	120	2 A33787	vascular endothel
2	45	81.8	133	2 B49530	vascular endothel
3	45	81.8	146	2 S57956	ovine vascular end
4	45	81.8	190	2 B40080	vascular endothel
5	42	76.4	190	2 B44881	vascular endothel
6	42	76.4	190	2 A35987	glioma-derived vas
7	42	76.4	214	2 A44881	vascular endothel
8	41	74.5	190	2 S52130	vascular endothel
9	41	74.5	232	2 A41551	vascular endothel
10	40	72.7	82	2 E36787	hypothetical prote
11	39	70.9	162	2 T01973	hypothetical prote
12	39	70.9	419	2 S69207	vascular endothel
13	37	67.3	92	1 NFHO2	neurophysin 2 - ho
14	37	67.3	107	1 NVMH2F	vasopressin / neur
15	37	67.3	144	2 A29101	vasopressin / neur
16	37	67.3	147	1 NVSH2	vasopressin / neur
17	37	67.3	164	1 NVHU2	vasopressin / neur
18	37	67.3	166	1 NVBO2	vasopressin / neur
19	37	67.3	166	1 NVPG2	vasopressin / neur
20	36	65.5	73	2 P85358	hypothetical prote
21	36	65.5	74	2 G85358	stress responsive
22	36	65.5	101	2 S07514	gene 5.3 protein -
23	36	65.5	149	2 A41236	placental growth f
24	36	65.5	261	2 D90123	hypothetical prote
25	36	65.5	328	2 AB0021	lacI-family transc
26	36	65.5	818	2 F96586	hypothetical prote
27	36	65.5	1115	2 A45761	Ca2+-transporting
28	36	65.5	3655	2 T38084	TRAP-like protein
29	35	63.6	128	2 I51295	vascular endothel

30	35	63.6	148	2 D49530	16k vascular endot
31	35	63.6	158	2 A56125	placental growth f
32	35	63.6	228	2 T39386	probable GI-S-spec
33	35	63.6	354	2 S59521	mucin 5ac - mouse
34	35	63.6	416	2 G86232	cysteine proteinas
35	35	63.6	477	2 S53362	mucin 5Ac (clone J
36	35	63.6	1058	2 JC1254	ubiquitin-protein
37	35	63.6	1058	2 A38564	ubiquitin-protein
38	35	63.6	4427	2 PN0637	polyketide synthas
39	35	63.6	351	2 T42421	hypothetical prote
40	34.5	62.7	1778	2 T30074	probable nucleopor
41	34	61.8	129	2 AD0255	probable phage ant
42	34	61.8	166	2 JN0248	platelet-derived g
43	34	61.8	196	2 A37359	platelet-derived g
44	34	61.8	196	2 B28964	platelet-derived g
45	34	61.8	196	2 A48851	platelet-derived g

ALIGNMENTS

RESULT 1
A33787
vascular endothelial growth factor (version 1) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 05-Nov-1999
C:Accession: A33787
R:Tischer, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.
Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989
A:Title: Vascular endothelial growth factor: a new member of the platelet-derived
A:Accession: A33787; MUID:90121225
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <TIS>
A:References: GB:M3750; NID:q163810; PIDN:AAA30805.1; PID:q163811
C:Keywords: alternative splicing

Query Match 81.8%; Score 45; DB 2; Length 120;
Best Local Similarity 77.8%; Pred. No. 1.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
DB 59 CCNESLIC 67

RESULT 2
B49530
vascular endothelial growth factor homolog A2R, 14.7K - Orf virus
C:Species: Orf virus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: B49530
R:Lytle, D.J.; Fraser, K.M.; Fleming, S.B.; Mercer, A.A.; Robinson, A.J.
J. Virol. 68, 84-92, 1994
A:Title: Homologs of vascular endothelial growth factor are encoded by the poxvir
A:Reference number: A49530; MUID:94076465
A:Contents: NZ2
A:Accession: B49530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <LYT>
A:Cross-references: GB:S67520; NID:q456897; PIDN:AA829220.1; PID:q456899
A>Note: sequence inconsistent with nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:141420, NCBIIP:141425)

Query Match 81.8%; Score 45; DB 2; Length 133;
Best Local Similarity 77.8%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESLIC 9

Db 70 CCNDESLC 78

RESULT 3

SV57956
Ovine vascular endothelial growth factor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C:Accession: S57956
R:Reimer, D.A.; Dal, Y.; Li, J.; Jones, S.C.; Moor, R.M.
Submitted to the EMBL Data Library, July 1995
A:Reference number: S57956
A:Accession: S57956
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RED>
A:Cross-references: EMBL:X89506; NID:g889350; PIDN:CAA61677.1; PID:g889351

Query Match

Best Local Similarity 81.8%; Score 45; DB 2; Length 146;
Pred. No. 1.6;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNDESLC 9
Db 85 CCNDESLC 93

RESULT 4

B40080
Vascular endothelial growth factor precursor (version 2) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Nov-1999
C:Accession: B40080; B33787; A33255
R:Leung, D.W.; Cachianes, G.; Kang, W.J.; Goeddel, D.V.; Ferrara, N.
Science 246, 1306-1309, 1989
A:Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.
A:Reference number: A40080; MUID:90069608
A:Accession: B40080
A:Molecule type: mRNA
A:Residues: 1-190 <LEU>
A:Cross-references: GB:M32976; NID:g163006; PIDN:AAA30502.1; PID:g163007
R:Risch, E.; Gospodarowicz, D.; Mitchell, R.; Silva, M.; Schilling, J.; Lau, K.; Crist
Biochem. Biophys. Res. Commun. 165, 1198-1206, 1989
A:Title: Vascular endothelial growth factor: a new member of the platelet-derived growth
A:Accession: B33787; MUID:90121225
A:Reference number: A33787
A:Molecule type: mRNA
A:Residues: 27-190 <TIS>
A:Cross-references: GB:M31836; NID:g163808; PIDN:AAA30804.1; PID:g163809
R:Ferrara, N.; Henzel, W.J.
Biochem. Biophys. Res. Commun. 161, 851-858, 1989
A:Title: Placental follicular cells secrete a novel heparin-binding growth factor specific
A:Reference number: A33255; MUID:89286596
A:Accession: A33255
A:Molecule type: protein
A:Residues: 27-31 <FER>
C:Keywords: alternative splicing; glycoprotein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-190/Product: vascular endothelial growth factor #status predicted <MAT>
F:100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 81.8%; Score 45; DB 2; Length 190;
Pred. No. 2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNDESLC 9
Db 85 CCNDESLC 93

RESULT 5

B44881
Vascular endothelial growth factor-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 05-Nov-1999
C:Accession: B44881; A43351; A61029
R:Breier, G.; Albrecht, U.; Sterrer, S.; Risau, W.
Development 114, 521-532, 1992
A:Title: Expression of vascular endothelial growth factor during embryonic angiogenesis
A:Reference number: A44881; MUID:92274860
A:Accession: B44881
A:Molecule type: mRNA
A:Residues: 1-190 <BRE>
A:Cross-references: GB:S38083; NID:g249858; PIDN:AA82253.1; PID:g249859
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:107622, NCBI:107623)
R:Claffey, K.P.; Wilkison, W.O.; Spiegelman, B.M.
J. Biol. Chem. 267, 16317-16322, 1992
A:Title: Vascular endothelial growth factor. Regulation by cell differentiation and
A:Reference number: A43351; MUID:92355593
A:Accession: A43351
A:Molecule type: mRNA
A:Residues: 1-116, 119-190 <CTA>
A:Cross-references: GB:M95200; NID:g202350; PIDN:AAA40547.1; PID:g202351
A:Note: sequence extracted from NCBI backbone (NCBIN:110665, NCBI:110675)
R:Rosenthal, R.A.; Megyesi, J.F.; Henzel, W.J.; Ferrara, N.; Folkman, J.
Growth Factors 4, 53-59, 1990
A:Title: Conditioned medium from mouse sarcoma 180 cells contains vascular endothelial
A:Reference number: A61029; MUID:91197543
A:Accession: A61029
A:Molecule type: protein
A:Residues: 27-38 <ROS>
C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; glycoprotein

Query Match

Best Local Similarity 76.4%; Score 42; DB 2; Length 190;
Pred. No. 6.3;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNDESLC 9
Db 85 CCNDESLC 93

RESULT 6

A35987
Glioma-derived vascular endothelial cell growth factor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 05-Nov-1999
C:Accession: A35987
R:Conn, G.; Bayne, M.L.; Soderman, D.D.; Kwok, P.W.; Sullivan, K.A.; Pallis, T.M.;
Proc. Natl. Acad. Sci. U.S.A. 87, 2628-2632, 1990
A:Title: Amino acid and cDNA sequences of a vascular endothelial cell mitogen that
A:Reference number: A35987; MUID:90207249
A:Accession: A35987
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-190 <CON>
A:Cross-references: GB:M32167; NID:g204287; PIDN:AAA41211.1; PID:g204288

Query Match

Best Local Similarity 76.4%; Score 42; DB 2; Length 190;
Pred. No. 6.3;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNDESLC 9
Db 85 CCNDESLC 93

RESULT 7

A44881
Vascular endothelial growth factor-3 precursor - mouse

N:Containing: vascular endothelial growth factor-2; vascular permeability factor
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 08-Oct-1999
C:Accession: A44881; C44881; A60932; S52136
R:Breier, G.; Albrecht, U.; Sterrer, S.; Risau, W.
Development 114, 521-532, 1992
A>Title: Expression of vascular endothelial growth factor during embryonic angiogenesis
A:Reference number: A44881; MUID:92274860
A:Accession: A44881
A:Molecule type: mRNA
A:Residues: 1-214 <BRE>
A:Cross-references: GB:S37052; NID:g249856; PIDN:AA62252.1; PID:g249857
A:Experimental source: embryo
A>Note: sequence extracted from NCBI backbone (NCBIN:104677, NCBI:104678)
A:Accession: C44881
A:Molecule type: mRNA
A:Residues: 1-140,209-214 <BR2>
A:Cross-references: GB:S38100; NID:g249860; PIDN:AA62254.1; PID:g249861
A>Note: sequence extracted from NCBI backbone (NCBIN:107624, NCBI:107625)
R:Claus, M.; Gerlach, M.; Gerlach, H.; Brett, J.; Wang, F.; Faller, P.C.; Pan, Y.C.
J. Exp. Med. 172, 1535-1545, 1990
A>Title: Vascular permeability factor: a tumor-derived polypeptide that induces endothelial
A:Reference number: A60932; MUID:91079755
A:Accession: A60932
A:Molecule type: protein
A:Residues: 27-33 <CLA>
R:Sugihara, T.; Kaul, S.C.; Mitsui, Y.; Madhwa, R.
Biochim. Biophys. Acta 1224, 365-370, 1994
A>Title: Enhanced expression of multiple forms of VEGF is associated with spontaneous im
A:Reference number: S52136; MUID:95101726
A:Accession: S52136
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-46 <SUG>
C:Comment: Homodimers could be demonstrated for recombinant VEGF-2 but not VEGF-3.
C:Keywords: alternative splicing; angiogenesis; disulfide bond; glycoprotein; homodimer;
F:1-26/Domain: signal sequence #status predicted <SUG>
F:27-214/Product: vascular endothelial growth factor-3 #status experimental <MAT>

Query Match 76.4%; Score 42; DB 2; Length 214;
Best Local Similarity 66.7%; Pred. No. 6.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CY 1 CCNEESLIC 9
111111111
DB 85 CCNDEALEC 93
RESULT 8
S52130
Vascular endothelial growth factor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C:Accession: S52130
R:Sharma, H.S.; Tang, Z.H.; Gho, B.C.G.; Verdouw, P.D.
Biochim. Biophys. Acta 1260, 235-238, 1995
A>Title: Nucleotide sequence and expression of the porcine vascular endothelial growth f
A:Reference number: S52130; MUID:95143284
A:Accession: S52130
A:Molecule type: preliminary
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-190 <SHA>
A:Cross-references: GB:X81380; NID:g587559; PIDN:CAA57143.1; PID:g587560

Query Match 74.5%; Score 41; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CY 1 CCNEESLIC 9
111111111
DB 85 CCNDEALEC 93

RESULT 9
A41551
Vascular endothelial growth factor 206 precursor - human
N:Alternate names: vascular permeability factor
N:Contents: vascular endothelial growth factor 121 (VEGF 121); VEGF 165; VEGF 189;
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
C:Accession: A41551; C41551; B41551; A40454; B40454; A40079; A40080; JQ146
R:Hock, K.A.; Ferrara, N.; Winer, J.; Cachianes, G.; Li, B.; Leung, D.W.
Mol. Endocrinol. 5, 1806-1814, 1991
A>Title: The vascular endothelial growth factor family: identification of a fourth
A:Reference number: A41551; MUID:92168017
A:Accession: A41551
A:Molecule type: mRNA
A:Residues: 1-232 <HOU1>
A:Cross-references: GB:S85192; NID:g246155; PID:g246156
A:Accession: C41551
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-140 'N', 183-232 <HOU2>
A:Accession: B41551
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-141,227-232 <HOU>
R:Tischer, E.; Mitchell, R.; Hartman, T.; Silva, M.; Gospodarowicz, D.; Fiddes, J.
J. Biol. Chem. 266, 11947-11954, 1991
A>Title: The human gene for vascular endothelial growth factor. Multiple protein
A:Reference number: A40454; MUID:91268072
A:Accession: A40454
A:Molecule type: DNA
A:Residues: 1-165,183-232 <TII>
A:Cross-references: GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63976
A:Accession: A40454
A:Molecule type: DNA
A:Residues: 1-140, 'N', 183-232 <TII2>
A:Cross-references: GB:M63971; GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63976
A:Accession: C40454
A:Molecule type: DNA
A:Residues: 1-141,227-232 <TII3>
A:Cross-references: GB:M63971; GB:M63972; GB:M63973; GB:M63974; GB:M63975; GB:M63976
R:Reck, P.J.; Hauser, S.D.; Krivt, G.; Sanzo, K.; Warren, T.; Feder, J.; Connolly
Science 246, 1309-1312, 1989
A>Title: Vascular permeability factor, an endothelial cell mitogen related to PDGF
A:Reference number: A40079; MUID:90069609
A:Accession: A40079
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-165,183-232 <KEC>
A:Cross-references: GB:M27281; NID:g340300; PIDN:AAA36807.1; PID:g340301
R:Leung, D.W.; Cachianes, G.; Kuang, W.J.; Goeddel, D.V.; Ferrara, N.
Science 246, 1306-1309, 1989
A>Title: Vascular endothelial growth factor is a secreted angiogenic mitogen.
A:Reference number: A40080; MUID:90069608
A:Accession: A40080
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <LEU>
A:Cross-references: GB:M32977; NID:g181970; PIDN:AAA35789.1; PID:g181971
R>Weinreb, K.; Marne, D.; Welch, H.A.
Biochem. Biophys. Res. Commun. 183, 1167-1174, 1992
A>Title: AIDS-associated Kaposi's sarcoma cells in culture express vascular endot
A:Reference number: JQ1463; MUID:92231879
A:Accession: JQ1463
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 183-232 <ME1>
A:Cross-references: EMBL:X62568; NID:g37658; PIDN:CAA44447.1; PID:g37659
A:Experimental source: AIDS-Kaposi's sarcoma cell
A:Accession: JQ1464
A:Molecule type: mRNA
A:Residues: 1-140, 'N', 227-232 <ME2>
A:Experimental source: AIDS-Kaposi's sarcoma cell

R:Connolly, D.T.; Olander, J.V.; Heuvelman, D.; Nelson, R.; Monsell, R.; Siegel, N.; Hay, J. Biol. Chem. 264, 20017-20024, 1989
 A:Title: Human vascular permeability factor. Isolation from U937 cells.
 A:Reference number: A34492; MUID:90062112
 A:Accession: A34492
 A:Molecule type: protein
 A:Residues: 27-36;43-49, 'R', 72-76, 'Q', 78-81;59-71 <CON>
 C:Comment: The most common of several alternatively spliced forms is VEGF 165.
 C:Genetics:
 A:Gene: GDB:VEGF
 A:Cross-references: GDB:132244; OMIM:192240
 A:Map position: 6p21-6p12
 C:Function:
 A:Description: promotes fluid and protein leakage from blood vessels.
 C:Keywords: alternative splicing; angiogenesis; dimer; disulfide bond; extracellular proc
 F:1-232/Product: vascular endothelial growth factor 206 precursor #status predicted <V20
 F:1-165,183-232/Product: vascular endothelial growth factor 189 precursor #status predi
 F:1-141,227-232/Product: vascular endothelial growth factor 121 precursor #status predi
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.5%; Score 41; DB 2; Length 232;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
 Db 86 CCNDEGLEC 94
 |||:|:|

RESULT 10
 E36787
 Hypothetical protein ORF13 - Ictalurid herpesvirus 1 (strain auburn 1)
 C:Species: Ictalurid herpesvirus 1
 A:Note: host Ictalurus punctatus (channel catfish)
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Oct-1999
 C:Accession: E36787
 R:Davidson, A.J.
 submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.
 A:Reference number: A36804
 A:Accession: E36787
 A:Molecule type: DNA
 A:Residues: 1-82 <DAV>
 A:Cross-references: GB:M75136; NID:9331209; PIDN:AAA88116.1; PID:9331223
 R:Davidson, A.J.
 Virology 186, 9-14, 1992
 A:Title: Channel catfish virus: a new type of herpesvirus.
 A:Reference number: A39447; MUID:92087490
 A:Contents: annotation
 A:Note: neither protein nor nucleic acid sequence is given
 C:Genetics:
 A:Gene: 13
 C:Keywords: DNA binding; zinc finger

Query Match 72.7%; Score 40; DB 2; Length 82;
 Best Local Similarity 66.7%; Pred. No. 7.1;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
 Db 11 CCNPMSLIC 19
 |||:|:|

RESULT 11
 T01973
 hypothetical protein T9A4.2 - Arabidopsis thaliana
 N:Alternate names: hypothetical protein F24G24.40
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 14-May-1999
 C:Accession: T01973; T04039

R:Zidanic, M.; McQuerry, Y.; Smith, A.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of A. thaliana T9A4.
 A:Reference number: Z14478
 A:Accession: T01973
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-162 <ZID>
 A:Cross-references: EMBL:AF096373; NID:93695400; PID:93695405
 A:Experimental source: cultivar Columbia
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, J.
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z15184
 A:Accession: T04039
 A:Molecule type: DNA
 A:Residues: 1-162 <BEV>
 A:Cross-references: EMBL:AL049488
 A:Experimental source: cultivar Columbia; BAC clone F24G24
 C:Genetics:
 A:Map position: 4
 A:Introns: 67/3
 A:Note: T9A4.2; F24G24.40

Query Match 70.9%; Score 39; DB 2; Length 162;
 Best Local Similarity 44.4%; Pred. No. 17;
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
 Db 17 CCSDAVALIC 25
 |||:|:|

RESULT 12
 S69207
 Vascular endothelial growth factor C precursor - human
 N:Alternate names: FLT4 ligand DHM
 C:Species: Homo sapiens (man)
 C:Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
 C:Accession: S69207; S61795; S71443; S69208; G02659
 R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Sak
 EMO J. 15, 1751, 1996
 A:Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a lig
 A:Reference number: S69207; MUID:96203094
 A:Accession: S69207
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-419 <JOU>
 A:Cross-references: EMBL:X94216; NID:91177488; PIDN:CAA63907.1; PID:6221096; PID:9
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1
 A:Note: only a part of the translation is shown
 A:Note: this is a revision to the sequence from reference S61795
 R:Joukov, V.; Pajusola, K.; Kaipainen, A.; Chillov, D.; Lahtinen, I.; Kukk, E.; Sak
 EMO J. 15, 290-298, 1996
 A:Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the F
 A:Reference number: S61795; MUID:96178224
 A:Accession: S61795
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 70-419 <JOU1>
 A:Note: this sequence has been revised in reference S69207
 A:Accession: S71443
 A:Molecule type: protein
 A:Residues: 'X', 104-120 <JOU2>
 R:Lee, J.; Gray, A.; Yuan, J.; Luo, S.M.; Avraham, H.; Wood, W.I.
 submitted to the EMBL Data Library, December 1995
 A:Description: Vascular endothelial growth factor related protein (VRP): A ligand
 A:Reference number: S69208
 A:Accession: S69208
 A:Molecule type: mRNA
 A:Residues: 1-419 <LEE>
 A:Cross-references: EMBL:U43142; NID:91150988; PIDN:AAA85214.1; PID:91150989
 R:Morris, J.C.

submitted to the EMBL Data Library, May 1996

A:Reference number: H01557

A:Accession: G02659

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-419 <MOR>

C:Genetics: EMBL:U06111; MTD:q1373426; PIDN:AAB02909.1; PID:q1373427

A:Gene: GDB:VEGFC; VBP

A:Cross-references: GDB:3890883; OMIM:601528

F:1-12/Domain: signal sequence #status predicted <SIG>

F:13-102/Domain: propeptide #status predicted <PRO>

F:103-419/Product: vascular endothelial growth factor C #status experimental <MAT>

Query Match 70.9%; Score 39; DB 2; Length 419;

Best Local Similarity 66.7%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9

Db 165 CCNEESLIC 173

RESULT 13

NF02

neurophysin 2 - horse (tentative sequence) (fragment)

C:Species: Equus caballus (domestic horse)

C>Date: 22-May-1981 #sequence_revision 22-May-1981 #text_change 31-Mar-2000

C:Accession: A01444

R:Chauvet, M.T.; Codogno, P.; Chauvet, J.; Acher, R.

FEBS Lett. 80, 374-376, 1977

A:Title: Phylogeny of the neurophysins: complete amino acid sequence of horse MSEL-neuro

A:Reference number: A01444; MUID:77246829

A:Accession: A01444

A:Molecule type: protein

A:Residues: 1-92 <CH>

C:Superfamily: oxytocin-neurophysin

C:Keywords: hormone; hypothalamus

F:7-51,10-24,18-41,25-31,58-70,64-82,71-76/Disulfide bonds: #status predicted

Query Match 67.3%; Score 37; DB 1; Length 92;

Best Local Similarity 62.5%; Pred. No. 24;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNEESLI 8

Db 70 CCNEESLI 77

RESULT 14

NW02F

vasopressin / neurophysin 2 precursor - finback whale (tentative sequence) (fragment)

C:Species: Balaenoptera physalus (finback whale, common rorqual)

C>Date: 18-Aug-1982 #sequence_revision 05-Jan-1996 #text_change 31-Mar-2000

C:Accession: B93147; A91447; A01446

R:Acher, R.; Chauvet, J.; Chauvet, M.T.

Nature 201, 191-192, 1964

A:Title: Isolation of finback whale oxytocin and vasopressin.

A:Reference number: A93147

A:Accession: B93147

A:Molecule type: protein

A:Residues: 1-9 <ACH>

R:Chauvet, M.T.; Codogno, P.; Chauvet, J.; Acher, R.

FEBS Lett. 88, 91-93, 1978

A:Title: Phylogeny of neurophysins. Complete amino acid sequence of whale (Balaenoptera

A:Accession: A91447

A:Reference number: A91447; MUID:78148313

A:Molecule type: protein

A:Residues: 13-107 <CHA>

A:Note: some residues may have been positioned only by homology with other neurophysins

C:Comment: Glycine is placed at 10 and X's at 11 and 12 by homology with the comp

C:Function:

A:Description: vasopressin is a hypothalamic peptide hormone that is an antidiuret

in for vasopressin

C:Superfamily: oxytocin-neurophysin

C:Keywords: amidated carboxyl end; glycoprotein; hormone; hypothalamus

F:1-9/Product: Arg-vasopressin #status experimental <VAS>

F:13-107/Product: neurophysin 2 #status experimental <NF2>

F:1-6/Disulfide bonds: #status experimental

F:9/Modified site: amidated carboxyl end (Gly) (amide in mature form from followi

F:22-66,25-39,33-56,40-46,73-85,79-97,86-91/Disulfide bonds: #status predicted

Query Match 67.3%; Score 37; DB 1; Length 107;

Best Local Similarity 62.5%; Pred. No. 27;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNEESLI 8

Db 85 CCNEESLI 92

RESULT 15

A29101

vasopressin / neurophysin 2 precursor - guinea pig

N:Alternate names: copeptin

C:Species: Arg-vasopressin; neurophysin 2; pituitary glycopeptide

C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 06-Sep-1996

C:Accession: A29101; S00009; J50300; A23630

R:Chauvet, M.T.; Rouille, Y.; Chauvet, J.; Acher, R.

FEBS Lett. 210, 40-44, 1987

A:Title: Guinea pig neurohypophysial hormones. Peculiar processing of the three-d

A:Reference number: A29101; MUID:87105929

A:Accession: A29101

A:Molecule type: protein

A:Residues: 1-144 <CHA>

R:Chauvet, J.; Chauvet, M.T.; Acher, R.

FEBS Lett. 217, 180-183, 1987

A:Title: Conformation limited proteolysis in the common neurophysin-copeptin precu

A:Reference number: S00009; MUID:87247214

A:Accession: S00009

A:Molecule type: protein

A:Residues: 13-144 <CH4>

R:Chauvet, M.T.; Chauvet, J.; Acher, R.

Int. J. Pept. Protein Res. 30, 676-682, 1987

A:Title: Guinea pig MSEL-neurophysin. Sequence comparison of eight mammalian MSEL-

A:Reference number: J50300; MUID:88138574

A:Accession: J50300

A:Molecule type: protein

A:Residues: 13-105 <CH2>

A:Note: 93-Ala was also found

R:Chauvet, M.T.; Chauvet, J.; Acher, R.

FEBS Lett. 197, 169-172, 1986

A:Title: Guinea pig copeptin. The glycopeptide domain of the vasopressin precursor

A:Reference number: A23630; MUID:86136563

A:Accession: A23630

A:Molecule type: protein

A:Residues: 107-144 <CH3>

C:Superfamily: oxytocin-neurophysin

C:Keywords: amidated carboxyl end; glycoprotein; hormone; hypothalamus

F:1-9/Product: Arg-vasopressin #status experimental <RVP>

F:13-105/Product: neurophysin 2 #status experimental <NF2>

F:1-6/Disulfide bonds: #status experimental <GCP>

F:9/Modified site: amidated carboxyl end (Gly) (amide in mature form from followi

F:22-66,25-39,33-56,40-46,73-85,79-97,86-91/Disulfide bonds: #status predicted

F:112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.3%; Score 37; DB 2; Length 144;

Best Local Similarity 62.5%; Pred. No. 34;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLI 8
11:11:
Db 85 CCNDESCV 92

Search completed: June 17, 2002, 15:58:23
Job time: 192 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:28 ; Search time 44.84 seconds
(without alignments)
7.772 Million cell updates/sec

Title: US-09-761-636A-9
Perfect score: 55
Sequence: 1 CCNESLIC 9

Scoring table: BLOSOM62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	354	1	VEGD_HUMAN
2	49	89.1	326	1	VEGD_RAT
3	45	81.8	133	1	VEGH_OREN2
4	45	81.8	146	1	VEGA_SHEEP
5	45	81.8	164	1	VEGA_CAVPO
6	45	81.8	190	1	VEGA_BOVIN
7	45	81.8	358	1	VEGD_MOUSE
8	42	76.4	214	1	VEGA_MOUSE
9	42	76.4	214	1	VEGA_RAT
10	41	74.5	190	1	VEGA_HORSE
11	41	74.5	190	1	VEGA_PIG
12	41	74.5	214	1	VEGA_CANFA
13	41	72.7	82	1	VG13_HSV1
14	40	72.7	126	1	VEGC_RAT
15	39	70.9	415	1	VEGC_MOUSE
16	39	70.9	419	1	VEGC_HUMAN
17	39	70.9	149	1	PIGF_HUMAN
18	38	69.1	92	1	NEU2_BOVIN
19	37	67.3	107	1	NEU2_HORSE
20	37	67.3	144	1	NEU2_BALPH
21	37	67.3	147	1	NEU2_CAVPO
22	37	67.3	164	1	NEU2_SHEEP
23	37	67.3	166	1	NEU2_HUMAN
24	37	67.3	166	1	NEU2_BOVIN
25	37	67.3	166	1	NEU2_PIG
26	37	67.3	190	1	VEGA_MESAU
27	36	65.5	73	1	RC23_ARATH
28	36	65.5	74	1	RC24_ARATH
29	36	65.5	101	1	Y53_BPT3
30	36	65.5	221	1	PIGF_HUMAN
31	36	65.5	3655	1	YAMB_SCHPO
32	35	63.6	148	1	VEGH_ORNP7
33	35	63.6	158	1	PIGF_RAT

ALIGNMENTS

RESULT ID	VEGD_HUMAN	STANDARD	PRT	354 AA.
AC	043915:			
DT	01-MAR-2002 (rel. 41, last sequence update)			
DT	01-MAR-2002 (rel. 41, last annotation update)			
DT	01-MAR-2002 (rel. 41, last annotation update)			
DE	Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).			
DE	FIGF OR VEGFD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=97349118; PubMed=9205122;			
RX	Yamada Y., Nezu J.-I., Shlman M., Hirata Y.;			
RT	"Molecular cloning of a novel vascular endothelial growth factor,			
RT	VEGF-D.";			
RL	Genomics 42:483-488(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=98140120; PubMed=9479493;			
RX	Rocchigliani M., Lestini M., Luddi A., Orlandini M., Franco B.,			
RA	Rossi E., Ballabio A., Zuffardi O., Oliviero S.;			
RT	"Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1			
RT	between the FIGA and the GRPR genes.";			
RL	Genomics 47:207-216(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9818549; PubMed=9435229;			
RA	Achen M.G., Jeltsch M., Kukk E., Maekinen T., Vitali A., Wilks A.F.,			
RT	Alitalo K., Stacker S.A.;			
RT	"Vascular endothelial growth factor D (VEGF-D) is a ligand for the			
RT	tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";			
RT	Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).			
RN	[4]			
RP	PROCESSING, AND SEQUENCE OF 89-94; 100-105 AND 206-213.			
RX	MEDLINE=20011413; PubMed=10542248;			
RA	Stacker S.A., Stenvers K.L., Caesar C., Vitali A., Domagala T.,			
RA	Nice E.C., Roufali S., Simpson R.J., Moritz R., Karpanen T.,			
RA	Alitalo K., Achen M.G.;			
RT	"Bioynthesis of vascular endothelial growth factor-D involves			
RT	proteolytic processing which generates non-covalent homodimers.";			
RL	J. Biol. Chem. 274:32127-32136(1999).			
CC	-I- FUNCTION: growth factor active in angiogenesis, lymphangiogenesis			
CC	and endothelial cell growth, stimulating their proliferation and			
CC	migration and also has effects on the permeability of blood			
CC	vessels. May function in the formation of the venous and lymphatic			
CC	vascular systems during embryogenesis, and also in the maintenance			
CC	of differentiated lymphatic endothelium in adults. Binds and			
CC	activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.			
CC	-I- SUBUNIT: Homodimer; non-covalent and antiparallel.			

34	35	63.6	158	1	PIGF_MOUSE	P49764	mus musculus
35	35	63.6	216	1	VEGA_CHICK	P52582	gallus galli
36	35	63.6	495	1	FASC_SMRPU	O05634	strongyloce
37	35	63.6	1058	1	UBA1_HUMAN	P22314	homo sapien
38	35	63.6	1058	1	UBA1_MOUSE	Q02053	mus musculus
39	35	63.6	1058	1	UBA1_RABIT	Q29504	oryctolagus
40	35	63.6	4427	1	PKSL_BACSU	O05470	baetilius su
41	34	61.8	204	1	PDGA_RAT	P28576	rattus norv
42	34	61.8	211	1	PDGA_HUMAN	P04085	homo sapien
43	34	61.8	211	1	PDGA_MOUSE	P20033	mus musculus
44	34	61.8	213	1	PDGA_RABIT	P24007	oryctolagus
45	34	61.8	225	1	CLDB_HUMAN	P56748	homo sapien

```

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in lung, heart, small
CC intestine and fetal lung, and at lower levels in skeletal muscle,
CC colon, and pancreas.
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
DR EMBL: D89630; BAA24264.1; -
DR EMBL: Y12863; CAA73370.1; -
DR EMBL: Y12864; CAA73371.1; -
DR EMBL: Y12865; CAA73371.1; JOINED.
DR EMBL: Y12866; CAA73371.1; JOINED.
DR EMBL: Y12867; CAA73371.1; JOINED.
DR EMBL: Y12868; CAA73371.1; JOINED.
DR EMBL: Y12869; CAA73371.1; JOINED.
DR EMBL: Y12870; CAA73371.1; JOINED.
DR EMBL: AJ000185; CAA03942.1; -
DR MIM: 300091; -
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KW Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 82 88
FT CHAIN 29 205
FT PROPEP 206 354
FT DOMAIN 222 318
FT REPEAT 222 237
FT REPEAT 258 273
FT REPEAT 277 293
FT REPEAT 301 318
FT DISULFID 111 153
FT DISULFID 142 189
FT DISULFID 146 191
FT DISULFID 136 136
FT DISULFID 145 145
FT CARBOHYD 155 155
FT CARBOHYD 185 185
FT CARBOHYD 287 287
SQ SEQUENCE 354 AA; 40444 MW; 2048D769D735173E CRC64;

Query Match 100.0%; Score 55; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

AC O35251;
DR 01-MAR-2002 (Rel. 41, Created)
DR 01-MAR-2002 (Rel. 41, Last sequence update)
DR 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
DE growth factor) (FIGF).
GN FIGF OR VEGFD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Yamada Y., Hirata Y., Nezu J., Shizume M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
CC and endothelial cell growth, stimulating their proliferation and
CC migration and also has effects on the permeability of blood
CC vessels. May function in the formation of the venous and lymphatic
CC vascular systems during embryogenesis, and also in the maintenance
CC of differentiated lymphatic endothelium in adults. Binds and
CC activates VEGFR-3 (Flt4) receptor (By similarity).
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
DR EMBL: AF014827; AAB66557.1; -
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PDGF; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS00278; PDGF_2; 1.
KW Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 325
FT DOMAIN 227 317
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 305 317
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 89.1%; Score 49; DB 1; Length 326;

```

Best Local Similarity 77.8%; Pred. No. 0.13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
Db 150 CCNESVMC 158

RESULT 3
VEGH_ORFN2 STANDARD; PRT; 133 AA.
AC P52584;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vascular endothelial growth factor precursor.
GN A2R.
OS Orf virus (strain NZ2) (OV NZ-2).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=10259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94076465; PubMed=8254780;
RA Lytle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
RT "Homologs of vascular endothelial growth factor are encoded by the
RT Poxvirus orf virus."
RL J. Virol. 68:84-92(1994).
CC -1- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S67520; AAB29220.2; -.
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 1 133
FT SIGNAL 21 133
FT CHAIN 21 133
FT DISULFID 36 78
FT DISULFID 67 112
FT DISULFID 71 114
FT DISULFID 61 61
FT DISULFID 70 70
FT CARBOHYD 85 85
SQ SEQUENCE 133 AA; 14715 MW; 917C0F688303C39 CRC64;

Query Match 81.8%; Score 45; DB 1; Length 133;
Best Local Similarity 77.8%; Pred. No. 0.29;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
Db 70 CCNDSLEEC 78

RESULT 4
VEGA_SHEEP STANDARD; PRT; 146 AA.
ID VEGA_SHEEP

AC P50412;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97117958; PubMed=8958842;
RA Redmer D.A., Dai Y., Li J., Charnock-Jones D.S., Smith S.K.,
RA Reynolds L.P., Moor R.M.;
RT "Characterization and expression of vascular endothelial growth
RT factor (VEGF) in the ovine corpus luteum."
RL J. Reprod. Fert. 108:157-165(1996).

CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X89506; CAM1677.1; -.
DR HSSP: P15692; 1VPP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Multigene family.
FT SIGNAL 1 26
FT CHAIN 1 146
FT DISULFID 27 146
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
SQ SEQUENCE 146 AA; 17247 MW; 4E792C8557E91760 CRC64;

Query Match 81.8%; Score 45; DB 1; Length 146;
Best Local Similarity 77.8%; Pred. No. 0.32;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
Db 85 CCNDSLEEC 93

RESULT 5
VEGA_CAVPO STANDARD; PRT; 164 AA.
ID VEGA_CAVPO
AC P26617;

01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Vascular endothelial growth factor A (VEGF-A) (Vascular permeability factor) (VPF).
VEGF OR VEGFA.
Cavia porcellus (Guinea pig).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
NCBI_TaxId=10141;
(1)
SEQUENCE FROM N.A.
Tissue=Blle duct;
Berser B.;
Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Growth factor active in angiogenesis, and endothelial cell growth. Induces endothelial proliferation and vascular permeability (By similarity).
-1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer with p16g (By similarity).
-1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or to the extracellular matrix unless released by heparin (By similarity).
-1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; M84230; AAA37057.1; -
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF_1.
DR ProDom; PD001629; PDGF_1.
DR SMART; SM00141; PDGF_1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR Mitogen; Angiogenesis; Growth factor; Glycoprotein.
KW DISULFID 25 67 BY SIMILARITY.
FT DISULFID 56 101 BY SIMILARITY.
FT DISULFID 60 103 BY SIMILARITY.
FT DISULFID 59 50 INTERCHAIN (BY SIMILARITY).
FT DISULFID 59 59 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 164 AA; 19330 MW; 9EB86A81A9D5DCA4 CRC64;

Query Match 81.8%; Score 45; DB 1; Length 164;
Best Local Similarity 77.8%; Pred. No. 0.35;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CNEESLIC 9
Db 59 CNEESLIC 67

RESULT 6
VEGA_BOVIN
ID VEGA_BOVIN STANDARD; PRT; 190 AA.
AC P15691;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular permeability factor) (VPF).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.
NCBI_TaxId=9913;
RX
RN
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
RX MEDLINE=90069608; PubMed=2479986;
RX Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic mitogen";
RL Science 246:1306-1309(1989).
RN
RP SEQUENCE OF 27-190 FROM N.A. (ISOFORMS ALPHA AND BETA).
RX MEDLINE=90121225; PubMed=2610687;
RX Tischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J.,
RA Lau K., Crisp T., Fiddes J.C., Abraham J.A.;
RT "Vascular endothelial growth factor: a new member of the platelet-derived growth factor gene family";
RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
RN
RP SEQUENCE OF 27-31.
RX MEDLINE=89286596; PubMed=2735925;
RA Ferrara N., Henzel W.J.;
RT "Follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells";
RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
CC
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and endothelial cell growth. It induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis, and induces permeabilization of blood vessels. It binds to the VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer with p16g (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or to the extracellular matrix unless released by heparin (By similarity).
CC
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: alpha (shown here) and beta; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; M32976; AAA30502.1; -
DR EMBL; M31836; AAA30804.1; -
DR EMBL; M33750; AAA30805.1; -
DR PIR; A33255; A33255.
DR PIR; A33787; A33787.
DR PIR; B40080; B40080.
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF_1.
DR ProDom; PD001629; PDGF_1.
DR SMART; SM00141; PDGF_1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal; Heparin-binding; Alternative splicing; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 139 183 MISSING (IN ISOFORM BETA).
FT VARSPLIC 184 184 R -> K (IN ISOFORM BETA).
SQ SEQUENCE 190 AA; 22310 MW; EDB9035E46E24789 CRC64;

Query Match 81.8%; Score 45; DB 1; Length 190;
 Best Local Similarity 77.8%; Pred. No. 0.4;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
 DB 85 CCNDESLIC 93

RESULT 7
 VEGF_MOUSE STANDARD: PRT: 358 AA.
 AC P97946;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (Figf).
 DE FIGF OR VEGFD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Flt1problast;
 RX MEDLINE=97030254; PubMed=8876195;
 RA Orlandini M., Marconcin L., Ferruzzi R., Oliviero S.;
 RT "Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
 RL [2]
 SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RX MEDLINE=97349118; PubMed=9205122;
 RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
 RT "Molecular cloning of a novel vascular endothelial growth factor, VEGF-D.";
 RL Genomics 42:483-488(1997).
 [3]
 DEVELOPMENTAL STAGE.
 RX MEDLINE=98288130; PubMed=9622638;
 RA Avenaghi V., Orlandini M., Acampora D., Oliviero S., Simeone A.;
 RT "Embryonic expression pattern of the murine figf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family.";
 RT Mech. Dev. 73:221-224(1998).
 RL [4]
 RECEPTOR SPECIFICITY.
 RX MEDLINE=21276411; PubMed=11279005;
 RA Baldwin M.E., Catmel B., Nice E.C., Roufail S., Hall N.E.,
 RA Stevens K.L., Karkkainen M.J., Alltalo K., Stacker S.A., Achen M.G.;
 RT "The specificity of receptor binding by vascular endothelial growth factor-d is different in mouse and man.";
 RL J. Biol. Chem. 276:19166-19171(2001).
 -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor.
 -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
 -1- SUBCELLULAR LOCATION: Secreted.
 -1- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.
 -1- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several body structures and organs of the embryo such as limb buds, acoustic ganglion, teeth, heart, anterior pituitary as well as lung and kidney mesenchyme, liver, derma, and perosteum of the vertebral column.

CC -1- INDUCTION: By the transcription factor c-fos.
 CC -1- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 CC bound by non-covalent interactions (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC -----
 CC EMBL: X99572; CAA67892.1; -
 CC EMBL: D89628; BAA14002.1; -
 CC HSSP: P15692; 1VP.
 CC MGD: MGI:108037; Figf.
 CC InterPro: IPR000072; PDGF.
 CC Pfam: PF00341; PDGF_1.
 CC ProDom: PD001629; PDGF_1.
 CC SMART: SM00141; PDGF_1.
 CC PROSITE: PS00249; PDGF_1; 1.
 CC PROSITE: PS00278; PDGF_2; 1.
 CC MitoGen: Growth factor; Glycoprotein; Signal; Repeat;
 CC Cleavage on pair of basic residues; Multigene family.
 KW SIGNAL 1 21
 KW PROPEP 22 93
 KW CHAIN 94 210
 KW PROPEP 211 358
 KW DOMAIN 227 323
 FT REPEAT 227 242
 FT REPEAT 263 278
 FT REPEAT 282 298
 FT REPEAT 306 323
 FT DISULFID 116 158
 FT DISULFID 147 194
 FT DISULFID 151 196
 FT DISULFID 141 141
 FT DISULFID 150 150
 FT DISULFID 160 160
 FT CARBOHYD 190 190
 FT CARBOHYD 292 292
 FT SEQUENCE 358 AA; 40908 MW; 6636B17FBF07037C CRC64;
 SQ
 Query Match 81.8%; Score 45; DB 1; Length 358;
 Best Local Similarity 66.7%; Pred. No. 0.69; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 1;
 OY 1 CCNESLIC 9
 DB 150 CCNDESLIC 158

[1]
 RN SEQUENCE FROM N.A. (ISOFORMS VEGF-1; VEGF-2 AND VEGF-3).
 RX MEDLINE=92274860; PubMed=1592003;
 RA Breier G., Albrecht U., Sterrer S., Risau W.;
 RT "Expression of vascular endothelial growth factor during embryonic
 RT angiogenesis and endothelial cell differentiation.";
 RT Development 114:521-532(1992).
 RN
 RP SEQUENCE FROM N.A. (ISOFORM VEGF-1).
 RX MEDLINE=92355593; PubMed=1644816;
 RA Claffey K.P., Wikstrom W.O., Spiegelman B.M.;
 RT "Vascular endothelial growth factor. Regulation by cell
 RT differentiation and activated second messenger pathways.";
 RL J. Biol. Chem. 267:16317-16322(1992).
 RN
 RP SEQUENCE OF 1-3 FROM N.A.
 RX MEDLINE=96216498; PubMed=8632007;
 RA Shima D.T., Kuroki M., Deutsch U., Ng Y., Adams A.P., D'Amore P.A.;
 RT "The mouse gene for vascular endothelial growth factor. Genomic
 RT structure, definition of the transcriptional unit, and
 RT characterization of transcriptional and post-transcriptional
 RT regulatory sequences.";
 RL J. Biol. Chem. 271:3877-3883(1996).
 CC
 -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
 CC endothelial cell growth. It induces endothelial cell
 CC proliferation, promotes cell migration, inhibits apoptosis, and
 CC induces permeabilization of blood vessels. It binds to the
 CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
 CC heparin (By similarity).
 CC
 -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
 CC with PlGF (By similarity).
 CC
 -1- SUBCELLULAR LOCATION: VEGF-1 and VEGF-2 are secreted while VEGF-3
 CC remains cell-surface associated unless released by heparin.
 CC
 -1- ALTERNATIVE PRODUCTS: 3 isoforms; VEGF-3/VEGF18 (shown here),
 CC VEGF-1/VEGF164 and VEGF-2/VEGF120; are produced by alternative
 CC splicing.
 CC
 -1- TISSUE SPECIFICITY: In developing embryos, expressed mainly in the
 CC choroid plexus, paraventricular neuroepithelium, placenta and
 CC kidney glomeruli. Also found in bronchial epithelium, adrenal
 CC gland and in seminiferous tubules of testis. High expression of
 CC VEGF continues in kidney glomeruli and choroid plexus in adults.
 CC
 -1- DOMAIN: VEGF-3 contains a basic insert which acts as a cell
 CC retention signal.
 CC
 -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL, S37052; AAB22252.1; -;
 DR EMBL, S38083; AAB22253.1; -;
 DR EMBL, S38100; AAB22254.1; -;
 DR EMBL, M95200; AAA40547.1; -;
 DR EMBL, U41383; -; NOT_ANNOTATED_CDS.
 DR PIR, A43351; A43351.
 DR HSSP, P15692; 2VPE.
 DR MGD, MGI:103178; Vegf.
 DR InterPro: IPR000072; PDGF.
 DR Pfam: PF00341; PDGF, 1.
 DR ProDom: PD001629; PDGF, 1.
 DR SMART, SM00141; PDGF, 1.
 DR PROSITE, PS00249; PDGF_1; 1.
 DR PROSITE, PS50278; PDGF_2; 1.
 KM Mitogen: Angiogenesis; Growth factor; Glycoprotein; Signal;
 KW Heparin-binding; Alternative splicing; Multigene family;
 FT SIGNAL 1
 FT CHAIN 27
 FT DISULFID 51
 FT 82 127
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 VASCULAR ENDOTHELIAL GROWTH FACTOR A.

FT DISULFID 86 129 BY SIMILARITY.
 FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT VARSPLIC 140 140 K -> N (IN ISOFORM VEGF-1).
 FT VARSPLIC 141 164 MISSING (IN ISOFORM VEGF-1).
 FT VARSPLIC 141 208 MISSING (IN ISOFORM VEGF-2).
 FT CONFLICT 117 118 GE -> ER (IN REF. 2).
 SQ SEQUENCE 214 AA; 25283 MW; B5540B51E4BBE17 CRC64;

Query Match 76.4%; Score 42; DB 1; Length 214;
 Best Local Similarity 66.7%; Pred. NO. 1.4;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESESLIC 9
 Db 85 CCNDEALEC 93

RESULT 9
 VEGA_RAT STANDARD; PRT; 214 AA.
 AC P16612; Q9QXG7; Q9QXG6; Q9JRX7;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
 DE permeability factor) (VPF).
 GN VEGF OR VEGFA.
 OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_Taxid=10116;
 RN
 RP SEQUENCE FROM N.A. (ISOFORM VEGF-A164), AND SEQUENCE OF 27-190.
 RX MEDLINE=90207249; PubMed=2320579;
 RA Conn G., Bayne M.L., Soderman D.D., Kwok P.W., Sullivan K.A.,
 RA Palisi T.M., Hope D.A., Thomas K.A.;
 RT "Amino acid and cDNA sequences of a vascular endothelial cell mitogen
 RT that is homologous to platelet-derived growth factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2628-2633(1990).
 RN
 RP SEQUENCE FROM N.A. (ISOFORMS VEGF-A188; VEGF-A164; VEGF-A144 AND
 RP VEGF-A120).
 RA Ishii H., Arkawa T., Okayama M., Oota I., Takuma T., Inomata K.;
 RT "Developmental expression of vascular endothelial growth factor-A
 RT (VEGF-A) splicing variants, VEGF-A188, VEGF-A164, and VEGF-A120 in rat
 RT mesenter muscle.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 27-40.
 RC TISSUE=Glial tumor;
 RX MEDLINE=95221439; PubMed=7706320;
 RA DiSalvo J., Bayne M.L., Conn G., Kwok P.W., Trivedi P.G.,
 RA Soderman D.D., Palisi T.M., Sullivan K.A., Thomas K.A.;
 RT "Purification and characterization of a naturally occurring vascular
 RT endothelial growth factor: placenta growth factor heterodimer.";
 RL J. Biol. Chem. 270:7717-7723(1995).
 CC
 -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
 CC endothelial cell growth. It induces endothelial cell
 CC proliferation, promotes cell migration, inhibits apoptosis, and
 CC induces permeabilization of blood vessels. It binds to the
 CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
 CC heparin (By similarity).
 CC
 -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
 CC with PlGF (By similarity).
 CC
 -1- SUBCELLULAR LOCATION: VEGF-A120 is acidic and freely secreted.
 CC VEGF-A164 is more basic, has heparin-binding properties and, although
 CC a significant proportion remains cell-associated, most is
 CC after secretion and is bound avidly by heparin and the
 CC extracellular matrix, although it may be released as a soluble

CC form by heparin, heparinase or plasmin (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: At least 4 isoforms: VEGF-A188 (shown here),
 CC VEGF-A164, VEGF-A144 and VEGF-A120; are produced by alternative
 CC splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in the pituitary, in brain, in
 CC particularly in supraoptic and paraventricular nuclei and the
 CC choroid plexus. Also found abundantly in the corpus luteum of
 CC the ovary and in kidney glomeruli.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M32167; AAA41211.1; -;
 CC EMBL; AF215725; AAF19211.1; -;
 CC EMBL; AF215726; AAF19212.1; -;
 CC EMBL; AF222779; AAF25958.1; -;
 CC PIR; A35987; A35987.
 CC HSSP; P15692; 1VPP.
 CC InterPro: IPR000072; PDGF.
 CC Pfam; PF00341; PDGF; 1.
 CC ProDom; PD001629; PDGF; 1.
 CC SMART; SM00141; PDGF; 1.
 CC DR PROSITE; PS00249; PDGF_1; 1.
 CC DR PROSITE; PS50278; PDGF_2; 1.
 CC DR Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
 CC Heparin-binding; Alternative splicing; Multigene family.
 CC KW
 CC FT CHAIN 1 26
 CC FT SIGNAL 1 26
 CC FT DISULFID 27 214 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
 CC FT DISULFID 51 93 BY SIMILARITY.
 CC FT DISULFID 82 127 BY SIMILARITY.
 CC FT DISULFID 86 129 BY SIMILARITY.
 CC FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 CC FT CARBOHYD 100 100 N-LINKED (GLCNAC...).
 CC FT VARSPLIC 140 140 K -> N (IN ISOFORM VEGF-A164).
 CC FT VARSPLIC 141 164 MISSING (IN ISOFORM VEGF-A164).
 CC FT VARSPLIC 141 208 MISSING (IN ISOFORM VEGF-A120).
 CC FT VARSPLIC 165 208 MISSING (IN ISOFORM VEGF-A144).
 CC FT CONFLICT 101 101 V -> A (IN REF. 2; AAF19212).
 CC SQ SEQUENCE 214 AA; 25239 MW; 60FBB876F5304946 CRC64;

Query Match 76.4%; Score 42; DB 1; Length 214;
 Best Local Similarity 66.7%; Pred. No. 1.4;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESSLIC 9
 DB 85 CCNDEALC 93

RESULT 10
 VEGA_HORSE
 ID VEGA_HORSE STANDARD; PRT; 190 AA.
 AC Q9GKRO;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
 DE permeability factor) (VPF).
 DE permealility factor) (VPF).
 GN VEGF OR VEGFA.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxId=9796;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Miura N., Misumi K., Kawahara K., Nakashima M., Fukumitsu S.,
 RA Kawabata H., Uto N., Oka T., Maruyama I., Sakamoto H.;
 RT "Cloning of cDNA and high-level expression of equine vascular
 RT endothelial growth factor (VEGF).";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
 CC cell growth. Induces endothelial proliferation and vascular
 CC permeability (By similarity).
 CC -1- SUBUNIT: Homodimer; disulfide-linked.
 CC with PLGF (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
 CC to the extracellular matrix unless released by heparin (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB053350; BAB20890.1; -;
 CC InterPro: IPR000072; PDGF.
 CC Pfam; PF00341; PDGF; 1.
 CC ProDom; PD001629; PDGF; 1.
 CC SMART; SM00141; PDGF; 1.
 CC DR PROSITE; PS00249; PDGF_1; 1.
 CC DR PROSITE; PS50278; PDGF_2; 1.
 CC DR Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
 CC Multigene family.
 CC KW
 CC FT CHAIN 1 26
 CC FT SIGNAL 1 26
 CC FT DISULFID 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
 CC FT DISULFID 51 93 BY SIMILARITY.
 CC FT DISULFID 82 127 BY SIMILARITY.
 CC FT DISULFID 86 129 BY SIMILARITY.
 CC FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
 CC FT CARBOHYD 100 100 N-LINKED (GLCNAC...). (POTENTIAL).
 CC SQ SEQUENCE 190 AA; 22312 MW; 87E9E161439E5F87 CRC64;

Query Match 74.5%; Score 41; DB 1; Length 190;
 Best Local Similarity 66.7%; Pred. No. 1.9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCNESSLIC 9
 DB 85 CCNDEGLC 93

RESULT 11
 VEGA_PIG
 ID VEGA_PIG STANDARD; PRT; 190 AA.
 AC P49151; Q9GL52;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
 DE permeability factor) (VPF).
 DE permealility factor) (VPF).
 GN VEGF OR VEGFA.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Heart;
 RX MEDLINE=95143284; PubMed=7841203;
 RA Sharma H.S., Tang Z.H., Cho B.C.H., Verdouw P.D.;
 RT "Nucleotide sequence and expression of the porcine vascular
 RT endothelial growth factor.";

```

RL Blochim. Biophys. Acta 1260:235-238(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee T., Canty J.M.;
RT "PCR cloning of porcine cardiac vascular endothelial growth factor
RT gene."
RL Submitted (NCV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X81380; CAA57143.1;
DR EMBL: AF185502; AAC33064.1;
DR HSSP: P15692; IVGH.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR Mitogen: Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 102 102 T -> A (IN REF. 2).
SQ SEQUENCE 190 AA; 22368 MW; 04D408BD7913047F CRC64;

```

```

Query Match 74.5%; Score 41; DB 1; Length 190;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CCNEESLIC 9
Db 85 CCNDEGLEC 93

```

```

RESULT 12
VEGA_CANFA STANDARD; PRT; 214 AA.
AC 09MYV3; Q9XSF5; Q9XSF4; Q9XSF3;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

```

```

OX NCBI_TaxId=9615;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM VEGF188).
RX MEDLINE-20125516; Pubmed-1061874;
RA Scheidegger P., Weiglhofer W., Suarez S., Kaser-Holtz B., Steiner R.,
RA Ballmer-Hofer K., Jaussi R.;
RT "Vascular endothelial growth factor (VEGF) and its receptors in tumor-
RT bearing dogs."
RL Biol. Chem. 380:1449-1454(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS VEGF188, VEGF-182 AND VEGF-164).
RA Tissue=Heart;
RC Jingjing L., Roque R.S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; VEGF188 (shown here),
CC VEGF182 and VEGF164; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ133758; CAB82426.1;
DR EMBL: AF133250; AAD29684.1;
DR EMBL: AF133249; AAD29683.1;
DR EMBL: AF133248; AAD29682.1;
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF_1.
DR ProDom: PD001629; PDGF_1.
DR SMART: SM00141; PDGF_1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS50278; PDGF_2; 1.
DR Mitogen: Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Alternative splicing; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 214
FT DISULFID 51 93 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPLIC 140 140 K -> N (IN ISOFORM VEGF-164).
FT VARSPLIC 141 141 MISSING (IN ISOFORM VEGF-164).
FT VARSPLIC 159 164 MISSING (IN ISOFORM VEGF-182).
FT CONFLICT 143 143 I -> V (IN REF. 2).
FT CONFLICT 161 161 P -> S (IN REF. 2).
SQ SEQUENCE 214 AA; 25175 MW; 0AC980A158C4AB27 CRC64;

```

```

Query Match 74.5%; Score 41; DB 1; Length 214;
Best Local Similarity 66.7%; Pred. No. 2.1;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CCNEESLIC 9
Db 85 CCNDEGLEC 93

```

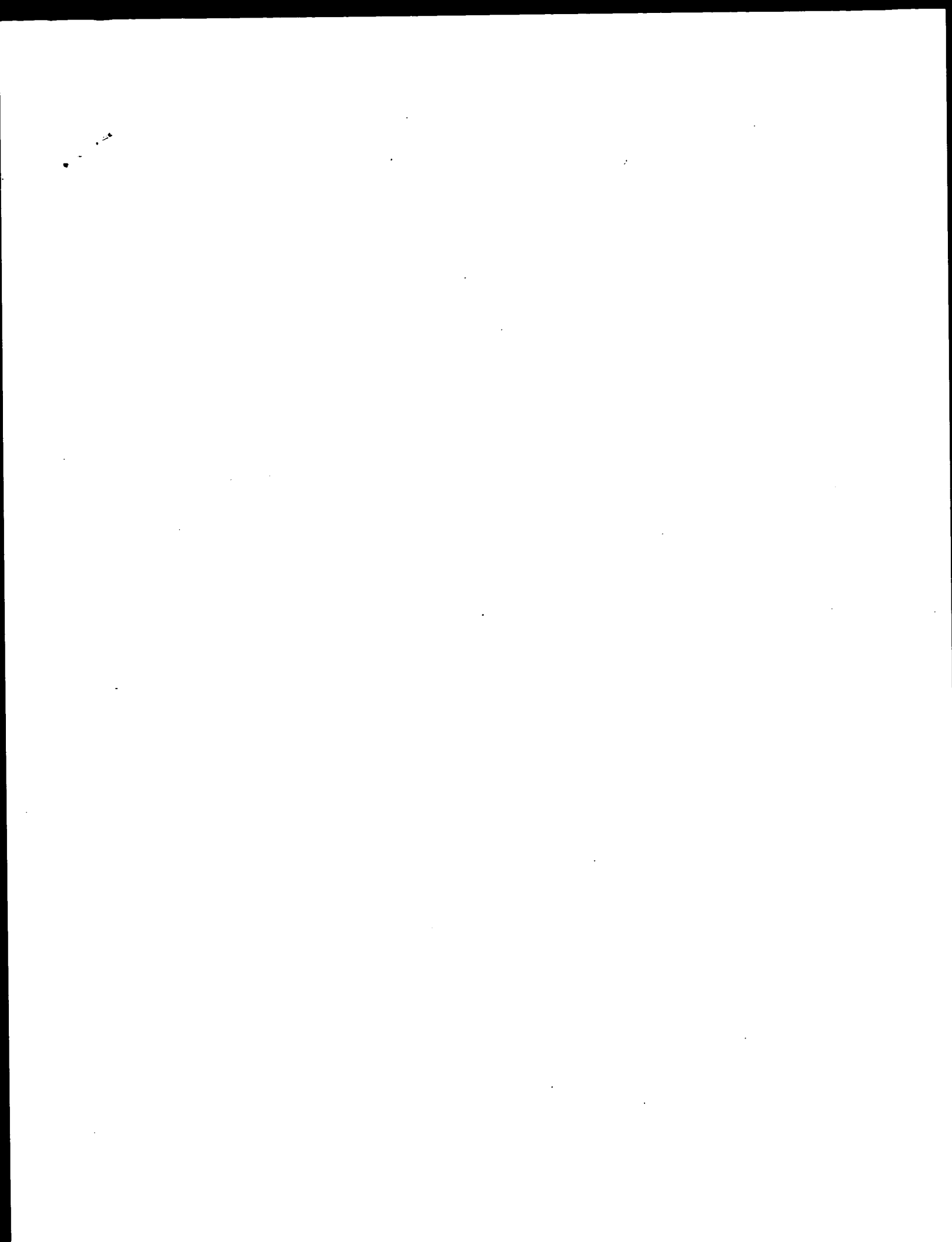
RESULT 13
VEGA_HUMAN STANDARD: PRT: 232 AA.
ID P15693; 016889; 060720; 075875; Q9U123; Q9UH58; Q9H1W9; Q9H1W8;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID:9606;
OX 11
RP SEQUENCE FROM N.A. (ISOFORM VEGF189 AND VEGF165).
RX MEDLINE-90069608; PubMed-2479986;
RA Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic mitogen.";
RL Science 246:1306-1309(1989).
RN 12
RP SEQUENCE FROM N.A. (ISOFORM VEGF189), AND PARTIAL SEQUENCE.
RX MEDLINE-90069609; PubMed-2479987;
RA Kerk P.J., Hauser S.D., Krivi G., Sanzo K., Warren T., Feder J., Connolly D.T.;
RT "Vascular permeability factor, an endothelial cell mitogen related to pgf.";
RL Science 246:1309-1312(1989).
RN 13
RP SEQUENCE FROM N.A. (ISOFORM VEGF189).
RX MEDLINE-91268072; PubMed-1711045;
RA Tischer E., Mitchell R., Hartman T., Silva M., Gospodarowicz D., Riddes J.C., Abraham J.A.;
RT "The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing.";
RL J. Biol. Chem. 266:11947-11954(1991).
RN 14
RP SEQUENCE FROM N.A. (ISOFORM VEGF206).
RX MEDLINE-92168017; PubMed-1791831;
RA Houck K.A., Ferrara N., Winer J., Cachianes G., Li B., Leung D.W.;
RT "The vascular endothelial growth factor family: identification of a fourth molecular species and characterization of alternative splicing of RNA.";
RL Mol. Endocrinol. 5:1806-1814(1991).
RN 15
RP SEQUENCE FROM N.A. (ISOFORM VEGF165).
RX MEDLINE-92231879; PubMed-1567395;
RA Weindel K., Marne D., Welch H.A.;
RT "AIDS-associated Kaposi's sarcoma cells in culture express vascular endothelial growth factor.";
RL Biochem. Biophys. Res. Commun. 183:1167-1174(1992).
RN 16
RP SEQUENCE FROM N.A. (ISOFORM VEGF145).
RX MEDLINE-97207275; PubMed-9054410;
RA Poltorak Z., Cohen T., Sivan R., Kandelis Y., Spira G., Vlodavsky I., Keshet E., Neufeld G.;
RT "VEGF145, a secreted vascular endothelial growth factor isoform that binds to extracellular matrix.";
RL J. Biol. Chem. 272:7151-7158(1997).
RN 17
RP SEQUENCE FROM N.A. (ISOFORM VEGF183).
RX TISSUE-Kidney;
RC MEDLINE-9906474; PubMed-9878851;
RA Lei J., Jiang A., Pei D.;
RT "Identification and characterization of a new splicing variant of vascular endothelial growth factor: VEGF183.";
RL Biochim. Biophys. Acta 1443:400-406(1998).
RN 18
RP SEQUENCE FROM N.A. (ISOFORM VEGF165).
RX TISSUE-Breast;
RC MEDLINE-98119755; PubMed-9450968;
RA Clafey K.P., Shih S.-C., Mullen A., Diennis S., Cusick J.L., Abrams K.R., Lee S.W., Detmar M.;
RT "Identification of a human VEGF/VEGF 3' untranslated region mediating hypoxia-induced mRNA stability.";
RL Mol. Biol. Cell 9:469-481(1998).
RN 19
RP SEQUENCE OF 114-209 FROM N.A. (ISOFORM VEGF183).
RX TISSUE-Retina;
RC MEDLINE-99165303; PubMed-10067980;
RA Jingling L., Xue Y., Agarwal N., Roque R.S.;
RT "Human Muller cells express VEGF183, a novel spliced variant of vascular endothelial growth factor.";
RL Invest. Ophthalmol. Vis. Sci. 40:752-759(1999).
RN 110
RP SEQUENCE FROM N.A. (ISOFORM VEGF165).
RX TISSUE-Hemangioma;
RC MEDLINE-99069608; PubMed-2479986;
RA Murata H., Fukushima J., Hattori S., Okuda K., Yanagi H.;
RT "Human CDNA for the vascular endothelial growth factor isoform VEGF165.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN 111
RP SEQUENCE FROM N.A. (ISOFORM VEGF148).
RX TISSUE-Renal glomerulus;
RC MEDLINE-99394945; PubMed-10464055;
RA Whittle C.J., Gillespie K.M., Harrison R., Mathieson P.W., Harper S.J.;
RT "Heterogeneous vascular endothelial growth factor (VEGF) isoform mRNA identification of VEGF148 mRNA, a novel truncated splice variant.";
RL Clin. Sci. 97:303-312(1999).
RN 112
RP SEQUENCE FROM N.A. (ISOFORM VEGF121).
RX Sato J.D., Whitney R.G.;
RT "Human CDNA for vascular endothelial growth factor isoform VEGF121.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN 113
RP SEQUENCE FROM N.A.
RX Williams S.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN 114
RP PRELIMINARY SEQUENCE OF 27-36; 43-50 AND 59-81.
RX MEDLINE-90062112; PubMed-2584205;
RA Connolly D.T., Olander J.V., Heuvelman D., Nelson R., Monseil R., Siegel N., Haymore B.L., Leingruber R., Feder J.;
RT "Human vascular permeability factor. Isolation from U937 cells.";
RL J. Biol. Chem. 264:20017-20024(1989).
RN 115
RP SEQUENCE OF 27-41.
RX MEDLINE-93145946; PubMed-7678805;
RA Fiebig B.L., Jaeger B., Schoellmann C., Weindel K., Wiltling J., Kochs G., Marne D., Hug H., Welch H.A.;
RT "Synthesis and assembly of functionally active human vascular endothelial growth factor homodimers in insect cells.";
RL Eur. J. Biochem. 211:19-26(1993).
RN 116
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 34-135.
RX MEDLINE-97352774; PubMed-9207067;
RA Muller Y.A., Li B., Christinger H.W., Wells J.A., Cunningham B.C., de Vos A.M.;
RT "Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:7192-7197(1997).
RN 117
RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS) OF 34-135.
RX MEDLINE-98035455; PubMed-9351807;
RA Muller Y.A., Christinger H.W., Keyt B.A., de Vos A.M.;
RT "The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93-A resolution: multiple copy flexibility and receptor binding.";
RL Structure 5:1325-1338(1997).
RN 118
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 39-134.
RX MEDLINE-99119204; PubMed-9922142;

RA Wiesmann C., Christinger H.W., Cochran A.G., Cunningham B.C.,
 RA Fairbrother W.J., Keenan C.J., Meng G., de Vos A.M.;
 RT "Crystal structure of the complex between VEGF and a receptor-binding
 RT peptide.";
 RL Biochemistry 37:17765-17772(1998).
 RN [19]
 RP STRUCTURE BY NMR OF 34-135.
 RX MEDLINE-97477915; PubMed-9336848;
 RA Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RT "1H, 13C, and 15N backbone assignment and secondary structure of the
 RT receptor-binding domain of vascular endothelial growth factor.";
 RL Protein Sci. 6:2250-2260(1997).
 RN [20]
 RP STRUCTURE BY NMR OF 137-215.
 RX MEDLINE-98298440; PubMed-9634701;
 RA Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A.,
 RA Starovasnik M.A.;
 RT "Solution structure of the heparin-binding domain of vascular
 RT endothelial growth factor.";
 RL Structure 6:637-648(1998).
 RN [21]
 RP FUNCTION.
 RX MEDLINE-21320570; PubMed-11427521;
 RA Murphy J.F., Fitzgerald D.J.;
 RT "Vascular endothelial growth factor induces cyclooxygenase-dependent
 RT proliferation of endothelial cells via the VEGF-2 receptor.";
 RL FASEB J. 15:1667-1669(2001).
 CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
 CC endothelial cell growth. It induces endothelial cell
 CC proliferation, promotes cell migration, inhibits apoptosis, and
 CC induces permeabilization of blood vessels. It binds to the
 CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
 CC heparin. Neuropilin-1 binds isoforms VEGF-165 and VEGF-145.
 CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
 CC with PlGF (by similarity).
 CC -1- SUBCELLULAR LOCATION: VEGFR121 is acidic and freely secreted.
 CC VEGFR165 is more basic, has heparin-binding properties and,
 CC although a significant proportion remains cell-associated, most is
 CC freely secreted. VEGFR189 is very basic; it is cell-associated
 CC after secretion and is bound avidly by heparin and the
 CC extracellular matrix, although it may be released as a soluble
 CC form by heparin, heparinase or plasmin.
 CC -1- ALTERNATIVE PRODUCTS: 7 isoforms: VEGFR206 (shown here), VEGFR189,
 CC VEGFR183, VEGFR165/VEGF, VEGFR148, VEGFR145 and VEGFR121; may be
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: The VEGFR189, VEGFR-165 and VEGFR-121 isoforms
 CC are widely expressed, whereas the VEGFR206 and VEGFR-145 are
 CC uncommon.
 CC -1- INDUCTION: Regulated by growth factors, cytokines, gonadotropins,
 CC nitric oxide, hypoxia, hypoglycemia and oncogenic mutations.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -1- DATABASE: NAME-R&D Systems' cytokine mini-reviews: VEGF;
 CC WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyid=230".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC -----
 CC Query Match 74.5%; Score 41; DB 1; Length 232;
 CC Best Local Similarity 66.7%; Pred. No. 2.3;
 CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC QY 1 CCNEESLIC 9
 CC Db 86 CCNDEGLEC 94
 CC -----
 CC RESULT 14
 CC VG13_HSV11 STANDARD; PRT; 82 AA.
 CC AC 000166;
 CC DT 01-DEC-1992 (Rel. 24, Created)
 CC DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical gene 13 zinc-binding protein.
 GN 13.
 OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC unclassified Herpesviridae.
 OX NCBI_TaxID=10401;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AUBURN 1;
 RX MEDLINE-92087490; PubMed-1727613;
 RA Davison A.J.;
 RT "Channel catfish virus: a new type of herpesvirus.";
 RL Virology 186:9-14(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC CC
 CC DR EMBL: M75136; AAA88194.1; -;
 CC DR EMBL: M75136; AAA88116.1; -;
 CC KW PIR: E36787; E36787.
 CC RN Hypothetical protein; Zinc; zinc-finger.
 CC SQ SEQUENCE 82 AA; 8821 MW; 6C009A50FF8C4C67 CRC64;
 CC -----
 CC Query Match 72.7%; Score 40; DB 1; Length 82;
 CC Best Local Similarity 66.7%; Pred. No. 1.4;
 CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC QY 1 CCNEESLIC 9
 CC Db 11 CCNPSMLIC 19
 CC -----
 CC RESULT 15
 CC VEGC_RAT STANDARD; PRT; 126 AA.
 CC AC 035737;
 CC DT 01-MAR-2002 (Rel. 41, Created)
 CC DT 01-MAR-2002 (Rel. 41, Last annotation update)
 CC DT 01-MAR-2002 (Rel. 41, Last annotation update)
 CC DE Vascular endothelial growth factor C precursor (VEGF-C) (Vascular
 CC endothelial growth factor related protein) (VNP) (Flt4 ligand) (Flt4-
 CC l) (Fragment).
 CC DE VEGFC.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Lung;
 RA Mandriota S.J., Pepper M.S.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
 CC cell growth, stimulating their proliferation and migration and
 CC also has effects on the permeability of blood vessels. May
 CC function in angiogenesis of the venous and lymphatic vascular
 CC systems during embryogenesis, and also in the maintenance of
 CC differentiated lymphatic endothelium in adults. Binds and
 CC activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
 CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
 CC variety of processed secreted forms with increased activity toward
 CC VEGFR-3, but only the fully processed form could activate VEGFR-2.
 CC VEGF-C first form an antiparallel homodimer linked by disulfide
 CC bonds. Before secretion, a cleavage occurs between arg-227 and

CC ser-228 producing an heterotetramer. The next extracellular step
 CC of the processing removes the N-terminal propeptide. Finally the
 CC mature VEGF-C is composed mostly of two VEGF homology domains
 CC (VHDS) bound by non-covalent interactions (By similarity).
 CC -1 SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF010302; AAB63248.1; -
 CC DR HSSP; P15692; 2VPF.
 CC DR InterPro; IPR000072; PDGF.
 CC DR Pfam; PF00341; PDGF; 1.
 CC DR ProDom; PD001629; PDGF; 1.
 CC DR SMART; SM00141; PDGF; 1.
 CC DR PROSITE; PS50278; PDGF_2; 1.
 CC KW Mitogen; Growth factor; Glycoprotein;
 CC KM Cleavage on pair of basic residues; Multigene family.
 CC FT NON_TER 1 1
 CC FT CHAIN 1 1
 CC FT PROPEP 72 71 VASCULAR ENDOTHELIAL GROWTH FACTOR C.
 CC FT DISULFID 6 53 POTENTIAL.
 CC FT DISULFID 10 55 BY SIMILARITY.
 CC FT DISULFID 9 9 INTERCHAIN (BY SIMILARITY).
 CC FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT NON_TER 126 126
 CC SQ SEQUENCE 126 AA; 13977 MM; 8F365AFBC4E037B0 CRC64;
 CC
 CC Query Match 70.9%; Score 39; DB 1; Length 126;
 CC Best Local Similarity 66.7%; Pred. No. 3;
 CC Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC QY 1 CCNEESLIC 9
 CC DB 9 CCNESEGLQC 17

Search completed: June 17, 2002, 16:24:28
 Job time: 1557 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:38 ; Search time 172.85 Seconds
(without alignments)
9.008 Million cell updates/sec

Title: US-09-761-636a-9
Perfect score: 55
Sequence: 1 CNEESLIC 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	89.1	326	091ZE4	091ZE4 rattus norv
2	45	81.8	68	097500	097500 oryctolagus
3	45	81.8	75	018843	018843 oryctolagus
4	45	81.8	78	09N1S2	09N1S2 capreolus c
5	45	81.8	118	09MZB1	09MZB1 ovis aries
6	45	81.8	123	09N1S1	09N1S1 capreolus c
7	45	81.8	132	09YMF3	09YMF3 orf virus
8	45	81.8	148	042571	042571 xenopus lae
9	45	81.8	190	077643	077643 ovis aries
10	45	81.8	194	042572	042572 xenopus lae
11	42	76.4	110	088911	088911 rattus norv
12	42	76.4	126	005458	005458 saccharomyc
13	42	76.4	141	070123	070123 mus musculu
14	42	76.4	144	073822	073822 brachydantio
15	42	76.4	188	073682	073682 brachydantio
16	42	76.4	190	09QX39	09QX39 spalax leuc

17	42	76.4	190	091ZE1	091ZE1 rattus norv
18	41	74.5	124	09GK00	09GK00 callithrix
19	41	74.5	126	09BDP7	09BDP7 macaca mula
20	41	74.5	169	096NW5	096NW5 homo sapien
21	41	74.5	189	095LQ4	095LQ4 felis silve
22	41	74.5	191	096L82	096L82 homo sapien
23	41	74.5	191	096KJ0	096KJ0 homo sapien
24	41	74.5	191	095NE5	095NE5 macaca fasc
25	40	72.7	895	09C1S1	09C1S1 candida gla
26	39	70.9	146	090X24	090X24 bothrops in
27	39	70.9	146	090X23	090X23 bothrops ja
28	39	70.9	162	082617	082617 arabidopsis
29	39	70.9	326	091ZM6	091ZM6 meriones un
30	39	70.9	415	091ZE3	091ZE3 rattus norv
31	39	70.9	418	057352	057352 coturnix co
32	38	69.1	411	09VCL7	09VCL7 drosophila
33	37	67.3	57	095S09	095S09 drosophila
34	37	67.3	142	09ERL6	09ERL6 mesocricetu
35	37	67.3	164	014935	014935 homo sapien
36	37	67.3	402	09UDM9	09UDM9 homo sapien
37	37	67.3	427	090586	090586 physarum po
38	37	67.3	458	060326	060326 homo sapien
39	37	67.3	620	09BQ18	09BQ18 homo sapien
40	36	65.5	110	097121	097121 bacterioph
41	36	65.5	261	098SA2	098SA2 guillardi
42	36	65.5	421	095T10	095T10 drosophila
43	36	65.5	502	09YEF4	09YEF4 drosophila
44	36	65.5	818	09SLI6	09SLI6 arabidopsis
45	36	65.5	943	09SRL7	09SRL7 arabidopsis

ALIGNMENTS

RESULT 1
ID 091ZE4 PRELIMINARY: PRT; 326 AA.
AC 091ZE4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VEGF-D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Kirkin V., Mazitschek R., Krishnan J., Steffen A., Waltenberger J.,
RA Pepper M.S., Giannits A., Sleeman J.P.;
RT "Characterization of Indolines which specifically inhibit VEGF-C-and
RT VEGF-D-induced activation of VEGFR-3 but not VEGFR-2.";
RL Eur. J. Biochem. 0:0-0(2001).
DR EMBL: AY032728; AAK96008.1;
SQ SEQUENCE 326 AA; 37106 MW; D7CAEBA6C9FAB7D CRC64;

Query Match 89.1%; Score 49; DB 11; Length 326;
Best Local Similarity 77.8%; Pred. No. 0.066;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNEESLIC 9
DB 150 CNEESVWC 158
RESULT 2
ID 097500 PRELIMINARY: PRT; 68 AA.
AC 097500;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Inoue K., Kawabe Y., Kodama T.;
RT "Rabbit VEGF cDNA, partial.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020216; BAA36949.1; -
DR HSSP; P15692; 2VPF.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF; 1.
DR ProDom; PD001629; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
FT NON_TER
FT 1
SQ SEQUENCE 68 AA; 7819 MW; 687638661E98DE0 CRC64;

Query Match 81.8%; Score 45; DB 6; Length 68;
Best Local Similarity 77.8%; Pred. No. 0.092;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
DB 40 CCNDESLC 48

RESULT 3
ID 018843 PRELIMINARY; PRT; 75 AA.
AC 018843;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NEW ZEALAND WHITE; TISSUE=SKLETAL MUSCLE;
RX MEDLINE=98191144; PubMed=9530113;
RA Skorjanc D., Jaschinski F., Heine G., Pette D.;
RT "Sequential increases in capillarization and mitochondrial enzymes in
RT low-frequency-stimulated rabbit muscle.";
RL Am. J. Physiol. 274:C810-C818(1998).
DR EMBL; AF022179; AAC15469.1; -
DR HSSP; P15692; 2VPF.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF; 1.
DR ProDom; PD001629; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
FT NON_TER
FT 1
SQ SEQUENCE 75 AA; 8720 MW; DDC2C5B29E69359 CRC64;

Query Match 81.8%; Score 45; DB 6; Length 75;
Best Local Similarity 77.8%; Pred. No. 0.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9

DB 28 CCNDESLC 36

RESULT 4
ID 09N1S2 PRELIMINARY; PRT; 78 AA.
AC 09N1S2;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR ISOFORM 121 (FRAGMENT).
GN VEGF.
OS Capreolus capreolus (Roe deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Odocoileinae; Capreolus.
OX NCBI_TaxID=9858;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=TESTIS;
RX MEDLINE=20532861; PubMed=11078967;
RA Wagener A., Blotner S., Goritz F., Fickel J.;
RT "Detection of growth factors in the testis of roe deer (Capreolus
RT capreolus).";
RL Anim. Reprod. Sci. 64:65-75(2000).
DR EMBL; AF152593; AAF73232.1; -
DR HSSP; P15692; 2VPF.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF; 1.
DR ProDom; PD001629; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
FT NON_TER
FT 1
SQ SEQUENCE 78 AA; 9131 MW; 7EE20DDEF17847C CRC64;

Query Match 81.8%; Score 45; DB 6; Length 78;
Best Local Similarity 77.8%; Pred. No. 0.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
DB 24 CCNDESLC 32

RESULT 5
ID 09M2B1 PRELIMINARY; PRT; 118 AA.
AC 09M2B1;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR (FRAGMENT).
GN VEGF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTAL ARTERY ENDOTHELIUM;
RA Zheng J., Tsai S.C., Magness R.R.;
RT "Growth factor expression in ovine fetal placental artery endothelial
RT cells.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF250375; AAF75258.1; -
DR HSSP; P15692; 1VPF.
DR InterPro; IPR000072; PDGF.
DR Pfam; PF00341; PDGF; 1.

DR PRODOM: PD001629; PDGF: 1.
DR SMART: SM00141; PDGF: 1.
DR PROSITE: PS00249; PDGF_1: 1.
DR PROSITE: PS50278; PDGF_2: 1.
FT NON_TER 1
SQ SEQUENCE 118 AA; 13931 MW; 757DC53AA56378A6 CRC64;

Query Match 81.8%; Score 45; DB 6; Length 118;
Best Local Similarity 77.8%; Pred. No. 0.16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 57 CCNDESLC 65

RESULT 6
ID 09N1S1 PRELIMINARY; PRT; 123 AA.
AC 09N1S1:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR ISOFORM 165 (FRAGMENT).
GN VEGF.
OS Capreolus capreolus (Roe deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Odocoileinae; Capreolus.
OX NCBI_TaxID=9858;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS:
RX MEDLINE=20532861; PubMed=11078967;
RA Wagener A., Blotner S., Goritz F., Fickel J.,
RT "Detection of growth factors in the testis of roe deer (Capreolus
RT capreolus).";
RL Anim. Reprod. Sci. 64:65-75(2000).
DR EMBL: AF152594; AAF73233.1; -
DR HSSP: P15692; 1VP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF: 1.
DR ProDom: PD001629; PDGF: 1.
DR SMART: SM00141; PDGF: 1.
DR PROSITE: PS00249; PDGF_1: 1.
DR PROSITE: PS50278; PDGF_2: 1.
FT NON_TER 1
FT NON_TER 123
SQ SEQUENCE 123 AA; 14354 MW; 0A756F54105AC61 CRC64;

Query Match 81.8%; Score 45; DB 6; Length 123;
Best Local Similarity 77.8%; Pred. No. 0.16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 24 CCNDESLC 32

RESULT 7
ID 09YMF3 PRELIMINARY; PRT; 132 AA.
AC 09YMF3:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR HOMOLOG VEGF-E.
GN VEGF-E.
OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=10258;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DL1701;
RX MEDLINE=99107753; PubMed=9889193;
RA Meyer M., Claus M., Lepple-Wienhues A., Waltenberger J.,
RA Augustin H.G., Ziche M., Lanz C., Buettnier M., Rzlin H.J., Dehio C.;
RT "A novel vascular endothelial growth factor encoded by orf virus,
RT VEGF-E, mediates angiogenesis via signalling through VEGFR-2 (KDR) but
RT not VEGFR-1 (Flt-1) receptor tyrosine kinases.";
RL EMO J. 18:363-374(1999).
DR EMBL: AF106020; AAD03735.1; -
DR HSSP: P15692; 1VP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF: 1.
DR ProDom: PD001629; PDGF: 1.
DR SMART: SM00141; PDGF: 1.
DR PROSITE: PS00249; PDGF_1: 1.
DR PROSITE: PS50278; PDGF_2: 1.
SQ SEQUENCE 132 AA; 14763 MW; 15F403A068B72926 CRC64;

Query Match 81.8%; Score 45; DB 12; Length 132;
Best Local Similarity 77.8%; Pred. No. 0.17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 69 CCNDESLC 77

RESULT 8
ID 042571 PRELIMINARY; PRT; 148 AA.
AC 042571:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR 122.
GN VEGF.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cleaver O., Tonissen K.F., Saha M.S., Krieg P.A.;
RT "Neovascularization of the xenopus embryo.";
RL Dev. Dyn. 0:0-0(1997).
DR EMBL: AF008593; AAB63679.1; -
DR HSSP: P15692; 1VP.
DR InterPro: IPR000072; PDGF.
DR Pfam: PF00341; PDGF: 1.
DR ProDom: PD001629; PDGF: 1.
DR SMART: SM00141; PDGF: 1.
DR PROSITE: PS00249; PDGF_1: 1.
DR PROSITE: PS50278; PDGF_2: 1.
SQ SEQUENCE 148 AA; 17234 MW; 4AD153CA2F8B1E95 CRC64;

Query Match 81.8%; Score 45; DB 13; Length 148;
Best Local Similarity 77.8%; Pred. No. 0.19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 86 CCNDESLC 94

RESULT 9
ID 077643 PRELIMINARY; PRT; 190 AA.
AC 077643:

DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR.
 GN VEGF.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA-RAMBOULLIET;
 RA Cheung C.Y., Brace R.A.;
 RT "Ovine vascular endothelial growth factor: Nucleotide sequence and
 RT expression in fetal tissues";
 RT Growth Factors 0:0-0(1998).
 RL EMBL; AF071015; AAC23608.1; -
 DR HSSP; P15692; 1VGH.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS0278; PDGF_2; 1.
 SQ SEQUENCE 190 AA; 22342 MW; 0D5E3B3E5C5E739 CRC64;

Query Match 81.8%; Score 45; DB 6; Length 190;
 Best Local Similarity 77.8%; Pred. No. 0.24;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNESLIC 9
 Db 85 CCNDESLC 93

RESULT 10
 ID 042572 PRELIMINARY; PRT; 194 AA.
 AC 042572;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR 196.
 GN VEGF.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=6355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cleaver O., Tonissen K.F., Saha M.S., Krieg P.A.;
 RT "Neovascularization of the xenopus embryo";
 RL Dev. Dyn. 0:0-0(1997).
 DR EMBL; AF008594; AAB63680.1; -
 DR HSSP; P15692; 1VGH.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS0278; PDGF_2; 1.
 SQ SEQUENCE 194 AA; 22672 MW; 85D7BEC7CEFEF17E CRC64;

Query Match 81.8%; Score 45; DB 13; Length 194;
 Best Local Similarity 77.8%; Pred. No. 0.25;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCNESLIC 9
 Db 85 CCNDESLC 93

Db 86 CCNDESLC 94

RESULT 11
 ID 088911 PRELIMINARY; PRT; 110 AA.
 AC 088911;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR A 110 (FRAGMENT).
 GN VEGF.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-PENIS;
 RX MEDLINE-99115228; PubMed-9916007;
 RA Burchard M., Burchard T., Chen M.W., Shabsigh A., de la Taille A.,
 RA Buttyan R., Shabsigh R.;
 RT "Expression of messenger ribonucleic acid splice variants for vascular
 RT endothelial growth factor in the penis of adult rats and humans";
 RL Biol. Reprod. 60:398-404(1999).
 DR EMBL; AF080594; AAC36708.1; -
 DR HSSP; P15692; 1VPP.
 DR InterPro: IPR000072; PDGF.
 DR Pfam; PF00341; PDGF; 1.
 DR ProDom; PD001629; PDGF; 1.
 DR SMART; SM00141; PDGF; 1.
 DR PROSITE; PS00249; PDGF_1; 1.
 DR PROSITE; PS0278; PDGF_2; 1.
 FT NON TER
 SQ SEQUENCE 110 AA; 12713 MW; B81B79AC08D89F06 CRC64;

Query Match 76.4%; Score 42; DB 11; Length 110;
 Best Local Similarity 66.7%; Pred. No. 0.57;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCNESLIC 9
 Db 59 CCNDESLC 67

RESULT 12
 ID 005458 PRELIMINARY; PRT; 126 AA.
 AC 005458;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE N2161 PROTEIN.
 GN N2161.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97245296; PubMed-9090055;
 RA De Antoni A., D'Angelo M., Dal Pero F., Sartorello F., Pandolfo D.,
 RA Pallavicini A., Lanfranchi G., Valle G.;
 RT "The DNA sequence of cosmid 14-11b from chromosome XIV of
 RT Saccharomyces cerevisiae reveals an unusually high number of
 RT overlapping open reading frames";
 RL Yeast 13:261-266(1997).
 DR EMBL; Z69382; CAA93403.1; -
 SQ SEQUENCE 126 AA; 12845 MW; 4FC6548A29E58628 CRC64;

Query Match 76.4%; Score 42; DB 3; Length 126;

Best local Similarity 66.7%; Pred. No. 0.65;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
DB 6 CCNDEALC 14

RESULT 13

070123 PRELIMINARY; PRT: 141 AA.

AC 070123; (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
VEGF15.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RC STRAIN-ICR;
RX MEDLINE=95101726; PubMed=7803491;
RA Sugihara T., Kaul S.C., Mitsui Y., Wadhwa R.;
RT "Enhanced expression of multiple forms of VEGF is associated with
spontaneous immortalization of murine fibroblasts.";
RL Biochim. Biophys. Acta 1224:365-370(1994).

RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-ICR;
RX MEDLINE=98112857; PubMed=9446618;
RA Sugihara T., Wadhwa R., Kaul S.C., Mitsui Y.;
RT "A novel alternatively spliced form of murine vascular endothelial
growth factor, VEGF 115.";
RL J. Biol. Chem. 273:3033-3038(1998).

DR EMBL; U50279; AAC05442.1; -.

DR HSSP; P15692; IVP.

DR MGD; MGI:103178; Vegf.

DR InterPro; IPR000072; PDGF.

DR Pfam; PF001629; PDGF_1.

DR PRODOM; PD001629; PDGF_1.

DR SMART; SM00141; PDGF_1.

DR PROSITE; PS00249; PDGF_1; 1.

DR PROSITE; PS0278; PDGF_2; 1.

SO SEQUENCE 141 AA; 15550 MW; A27CAEF5A7071338 CRC64;

Query Match 76.4%; Score 42; DB 11; Length 141;
Best Local Similarity 66.7%; Pred. No. 0.72;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
DB 85 CCNDEALC 93

RESULT 14

073822 PRELIMINARY; PRT: 144 AA.

AC 073822;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DE 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
VEGF.

OS Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

NCBI_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Liang D., Ge R.;

RT "Vascular endothelial growth factor 121 isoform from zebrafish, Danio
rerio.";

RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF059661; AAC14713.1; -.

DR HSSP; P15692; 2VPE.

DR InterPro; IPR000072; PDGF.

DR Pfam; PF00341; PDGF_1.

DR PRODOM; PD001629; PDGF_1.

DR SMART; SM00141; PDGF_1.

DR PROSITE; PS00249; PDGF_1; 1.

DR PROSITE; PS0278; PDGF_2; 1.

SO SEQUENCE 144 AA; 16479 MW; 303E6A7407AA0832 CRC64;

Query Match 76.4%; Score 42; DB 13; Length 144;
Best Local Similarity 66.7%; Pred. No. 0.73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
DB 83 CCNDEALC 91

RESULT 15

073682 PRELIMINARY; PRT: 188 AA.

AC 073682;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DE 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
VEGF.

OS Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Danio.

NCBI_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98222600; PubMed=9545518;
RA Liang D., Xu X., Chin A.J., Balasubramanian N.V., Teo M.A., Lam T.J.,
Weinberg E.S., Ge R.;

RT "Cloning and characterization of vascular endothelial growth factor
(VEGF) from zebrafish, Danio rerio.";

RL Biochim. Biophys. Acta 1397:14-20(1998).

DR EMBL; AF016244; AAC1274.1; -.

DR HSSP; P15692; IYGH.

DR ZFIN; ZDB-GENE-990415-273; vegf.

DR InterPro; IPR000072; PDGF.

DR Pfam; PF00341; PDGF_1.

DR PRODOM; PD001629; PDGF_1.

DR SMART; SM00141; PDGF_1.

DR PROSITE; PS00249; PDGF_1; 1.

DR PROSITE; PS0278; PDGF_2; 1.

SO SEQUENCE 188 AA; 21756 MW; FD203334EF365D1E CRC64;

Query Match 76.4%; Score 42; DB 13; Length 188;
Best Local Similarity 66.7%; Pred. No. 0.94;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
DB 83 CCNDEALC 91

Search completed: June 17, 2002, 16:23:39
Job time: 1613 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:12 ; Search time 227.35 Seconds
(without alignments)
4.397 Million cell updates/sec

Title:	US-09-761-636A-9
Perfect score:	55
Sequence:	1 CCNESLIC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      747574 segs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : A_Geneseq_032802:*

1	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1980.DAT.*
2	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1981.DAT.*
3	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1982.DAT.*
4	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1983.DAT.*
5	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1984.DAT.*
6	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1985.DAT.*
7	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1986.DAT.*
8	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1987.DAT.*
9	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1988.DAT.*
10	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1989.DAT.*
11	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1990.DAT.*
12	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1991.DAT.*
13	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1992.DAT.*
14	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1993.DAT.*
15	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1994.DAT.*
16	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1995.DAT.*
17	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1996.DAT.*
18	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1997.DAT.*
19	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1998.DAT.*
20	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA1999.DAT.*
21	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA2000.DAT.*
22	/SIDS1/gcgdata/hold_geneseq/genseqp_emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	55	100.0	9	22	AAU04528	VEGF based bicyclic
2	55	100.0	73	22	AAU04522	Human VEGF-D amino
3	55	100.0	96	22	AAU04520	Human VEGF-D amino
4	55	100.0	109	20	AAU23889	Human vascular endot
5	55	100.0	109	21	AAB11931	Human truncated VEG
6	55	100.0	325	19	AAW53240	Homo sapiens vascu
7	55	100.0	325	22	AAI97572	Homo VEGF-D prote
8	55	100.0	354	19	AAW49036	Human zvegf2 growt
9	55	100.0	354	19	AAW53241	Homo sapiens vascu
10	55	100.0	354	19	AAW44293	Human vascular end
11	55	100.0	354	21	AAB10649	Human VEGD protein

12	55	100.0	354	21	AAB32904.9	Human VEGF-D prote
13	55	100.0	354	21	AAV70750	Human prepro-vascu
14	55	100.0	354	21	AAV70963	Human vascular endo
15	100.0	354	22	AAU08441	Poly peptide for hu	
16	55	100.0	354	22	AAV97573	Human VEGF-D1 prot
17	55	100.0	354	22	AAB37606	Human VEGF-D. Hom
18	50	90.9	9	22	AAD04552	VEGF based blycili
19	49	89.1	9	22	AAU04553	VEGF based blycili
20	48	89.1	326	19	AAW44236	Rat vascular endot
21	48	87.3	354	22	AAB70685	Human vascular end
22	48	87.3	620	18	AAW14994	Human c-Fos induce
23	46	83.6	8	22	AAU04525	VEGF based monocy
24	45	81.8	101	20	AAM86229	Poxvirus ORF-1 tru
25	45	81.8	106	20	AAW86228	Poxvirus ORF-1 tru
26	45	81.8	111	20	AAW86227	Poxvirus ORF-1 tru
27	45	81.8	116	20	AAW86226	Poxvirus ORF-1 tru
28	45	81.8	120	12	AAU10916	Bovine vascular en
29	45	81.8	120	14	AAB38916	Bovine VEGF-120.
30	45	81.8	132	18	AAW40305	Parapox virus Hind
31	45	81.8	132	20	AAV34343	Parapox virus VEGF
32	45	81.8	132	21	AAV92776	Orf virus strain N
33	45	81.8	132	22	AAU08466	Poly peptide for or
34	45	81.8	132	22	AAV97574	Human VEGF-E prote
35	45	81.8	133	20	AAV34335	Parapox virus VEGF
36	45	81.8	133	20	AAV34443	Parapox virus rppv
37	45	81.8	133	20	AAW86208	Poxvirus ORF-1 pro
38	45	81.8	133	20	AAW86236	Poxvirus ORF-1 (VE
39	45	81.8	133	21	AAV92775	Orf virus strain N
40	45	81.8	146	20	AAV33438	Parapox virus VEGF
41	45	81.8	164	12	AAU10911	Bovine vascular en
42	45	81.8	164	14	AAB38920	Bovine VEGF-164.
43	45	81.8	178	20	AAV08287	Human growth facto
44	45	81.8	190	11	AAU08001	Bovine vascular en
45	45	81.8	321	19	AAW53243	Mus musculus vascu

ALIGNMENTS

RESULT	1
AAU04528	AAU04528 standard; Protein; 9 AA.
AAU04528;	
26-SEP-2001	(first entry)
VEGF based bicyclic dimeric peptide #2.	
Human: VEGF: vascular endothelial growth factor; angiogenesis; neovascularisation; lymphangiogenesis; psoriasis; tumour; diabetes induced neovascular sequelae; rheumatoid arthritis; diabetic retinopathy; chronic inflammation; cyclic.	
Synthetic.	
Key	Location/Qualifiers
Disulfide-bond	2..9
Disulfide-bond	/note- "This bond cyclises the peptide"
Disulfide-bond	1
/note- "A disulfide bond forms between residue 1 and residue 1 of an identical peptide to form a dimeric peptide, or between residue 1 and residue 17 of the sequence appearing as AAU04527 also forming a dimeric peptide"	
MO200152875-A1.	
26-JUL-2001.	
18-JAN-2001; 2001WO-US01533.	
18-JAN-2000; 2000US-0176293.	

W0200152875-A1.

26-JUL-2001.

18-JAN-2001; 2001WO-US01533.

18-JAN-2000; 2000US-0176293.

```

PR      16-MAY-2000; 2000US-0204590.
XX
XX PA      (LUDW-) LUDWIG INST CANCER RES.
XX
XX PI      Achen MG, Hughes RA, Stacker S, Cendron A;
XX WPI; 2001-442248/47.
XX
XX PT      Novel monomeric monocyclic peptide, used to interfere with
XX angio genesis, or lymphangiogenesis, is produced by cyclising a peptide
XX loop fragment from an exposed loop of a growth factor protein by
XX oxidizing the cysteine residues -
PS Claim 59; Page 32; 102pp; English.
CC
CC The sequence represents a dimeric bicyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides), and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised post-angioplasty restenosis, arthritis,
CC cerebrovascular accident, malignant or benign tumour, post-recovery
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX SQ
XX Sequence 9 AA:
XX
XX Query Match 100.0%; Score 55; DB 22; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 6.4e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CONESLIC 9
XX ||||||||
DB 1 coneslic 9
XX
XX RESULT 2
XX ID AU04522 standard; protein; 73 AA.
XX AU04522;
XX
XX 26-SEP-2001 (first entry)
XX
XX Human VEGF-D amino acids Val101-Thr 173.
DE
XX Human: VEGF-D; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation.
XX
XX Homo sapiens
OS
XX
XX WO200152875-A1.

```

XX		26-JUL-2001.
PD		
XX		
PF		18-JAN-2001; 2001WO-US01533.
XX		
PR		18-JAN-2000; 2000US-0176293.
XX		
PR		16-MAY-2000; 2000US-0204590.
XX		
PA		(LUDW-) LUDWIG INST CANCER RES.
XX		
PI	Achen MG, Hughes RA, Stacker S, Cendron A;	
DR	WPI; 2001-442248/47.	
XX		
PT	Novel monomeric monocyclic peptide, used to interfere with	
PT	angiogenesis or lymphangiogenesis, is produced by cyclising a peptide	
PT	loop fragment from an exposed loop of a growth factor protein by	
PT	oxidizing the cysteine residues -	
XX		
PS	Example 1; Page 90-91; 102pp; English.	
XX		
CC	The sequence represents Human VEGF-D (vascular endothelial growth factor)	
CC	amino acids Val101-Thr 173, used together with the C-terminal 23	
CC	residues of VEGF to make a hybrid theoretical molecule for 3	
CC	dimensional modelling. The sequence is used in a method of producing	
CC	a monomeric monocyclic peptide by a measuring beta-beta carbon separation	
CC	distances on opposite antiparallel strands of a peptide loop fragment	
CC	from an exposed loop of a growth factor protein and cyclising the peptide	
CC	by oxidising the cysteine residues. The monocyclic peptides' dimeric	
CC	bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic	
CC	peptide with at least one amino acid deleted prior to cyclisation are	
CC	used to interfere with angiogenesis, neovascularisation or	
CC	lymphangiogenesis in a mammal with a condition characterised by	
CC	angiogenesis, neovascularisation or lymphangiogenesis. The condition is	
CC	diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised	
CC	malignant or benign tumour, post-recovery cerebrovascular accident,	
CC	post-angioplasty restenosis, head, heat or cold trauma, substance-induced	
CC	neovascularisation of the liver, excessive hormone-related angiotenic	
CC	dysfunction, diabetes induced neovascular sequelae, hypertension induced	
CC	neovascular sequelae, or chronic liver infection. The peptides are also	
CC	used to modulate vascular permeability in a mammal (the mammal has a	
CC	condition characterised by fluid accumulation in peripheral limbs or in	
CC	lungs, peritoneal cavity, pleura, or brain. The peptides are used to	
CC	image blood vessels and lymphatic vasculature. The monomeric and bicyclic	
CC	peptides are used to interfere with at least one biological activity	
CC	induced by VEGF, VEGF-C or -D and are also used in combination with an	
CC	anti-inflammatory agent, to treat a chronic inflammation, especially	
CC	rheumatoid arthritis, psoriasis and diabetic retinopathy.	
XX		
SQ	Sequence 73 AA:	
	Query Match 100.0%; Score 55; DB 22; Length 73;	
	Best Local Similarity 100.0%; Prid. No. 0.22;	
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 CCNESLIC 9 	
Db	45 ccneeslic 53	
RESULT 3		
ID	AAU04520	
XX	AAU04520 standard; protein; 96 AA.	
AC	AAU04520;	
XX		
DT	26-SEP-2001 (first entry)	
DE	Human VEGF-D amino acids Val101-PRO186.	
XX		
KW	Human; VEGF-D; vascular endothelial growth factor; angiogenesis; neovascularisation; lymphangiogenesis; psoriasis; tumour;	

KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KM diabetic retinopathy; chronic inflammation.
XX
OS Homo sapiens.
XX
XX WO200152875-A1.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01533.
XX
XX 18-JAN-2000; 2000US-0176293.
XX
XX 16-MAY-2000; 2000US-0204590.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Hughes RA, Stacker S, Cendron A;
XX
XX WPI: 2001-442248/47.
XX
XX Novel monomeric monocyclic peptide, used to interfere with
XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
XX loop fragment from an exposed loop of a growth factor protein by
XX oxidizing the cysteine residues -
XX
XX Example 1; Page 89; 102pp; English.
XX
XX The sequence represents human VEGF-D (vascular endothelial growth factor)
XX amino acids Val101-PRO186. The sequence is used in a method of producing
XX a monomeric monocyclic peptide by a measuring beta-beta carbon separation
XX distances on opposite antiparallel strands of a peptide loop fragment
XX from an exposed loop of a growth factor protein and cyclising the peptide
XX by oxidising the cysteine residues. The monocyclic peptides, dimeric
XX bicyclic peptides (comprising 2 linked monocyclic peptides) and a cyclic
XX peptide with at least one amino acid deleted prior to cyclisation are
XX used to interfere with angiogenesis, neovascularisation or
XX lymphangiogenesis in a mammal with a condition characterised by
XX angiogenesis, neovascularisation or lymphangiogenesis. The condition is
XX diabetic retinopathy, psoriasis, arthropathy, hemangioma, vascularised
XX malignant or benign tumour, post-recovery cerebrovascular accident,
XX post-angioplasty restenosis, head, heat or cold trauma, substance-induced
XX neovascularisation of the liver, excessive hormone-related angiogenic
XX dysfunction, diabetes induced neovascular sequelae, hypertension induced
XX neovascular sequelae, or chronic liver infection. The peptides are also
XX used to modulate vascular permeability in a mammal (the mammal has a
XX condition characterised by fluid accumulation in peripheral limbs or in
XX lungs, peritoneal cavity, pleura, or brain. The peptides are used to
XX image blood vessels and lymphatic vasculature. The monomeric and bicyclic
XX peptides are used to interfere with at least one biological activity
XX induced by VEGF, VEGF-C or -D and are also used in combination with an
XX anti-inflammatory agent, to treat a chronic inflammation, especially
XX rheumatoid arthritis, psoriasis and diabetic retinopathy.
XX
SQ Sequence 96 AA;

Query Match 100.0%; Score 55; DB 22; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 45 ccneeslic 53
IIIIIIIIII

RESULT 4
AA123889
ID AA123889 standard; Protein: 109 AA.
XX
XX AA123889;
XX
XX 21-SEP-1999 (first entry)
XX

DE Human vascular endothelial growth factor (VEGF)-D.
XX
XX Vascular endothelial growth factor: VEGF; VEGF-D; malignant melanoma;
XX tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
XX wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
XX
XX Homo sapiens.
XX
XX WO9933485-A1.
XX
XX 08-JUL-1999.
XX
XX 23-DEC-1998; 98WO-US27373.
XX
XX 29-MAY-1998; 98US-0087392.
XX
XX 24-DEC-1997; 97AU-0001131.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Altalo K, Stacker SA;
XX
XX WPI: 1999-405368/34.
XX
XX A human cell line stably expressing vascular endothelial growth
XX factor D, useful for treating melanomas or tumours expressing VEGF-D
XX
XX Claim 6; Page 72; 79pp; English.
XX
XX The present sequence represents human vascular endothelial growth
XX factor (VEGF)-D. The specification describes a human cell line
XX which stably expresses VEGF-D, or fragments/analogues having VEGF-D
XX biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
XX or triplex DNA, VEGF-D variants or antibodies (especially chimeric
XX antibodies), are useful for the treatment or alleviation of malignant
XX melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
XX stimulating amounts of VEGF-D can be administered to enhance the
XX acceptance and/or healing of skin grafts or to stimulate the healing
XX of a surgical or traumatic wound to the skin. Lymphangiogenesis
XX stimulating amounts of VEGF-D can be used to treat lymphedema.
XX Endothelial proliferation stimulating amounts of VEGF-D are used to
XX treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
XX be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
XX useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
XX can be used to stimulate at least one VEGF-D bioactivity chosen from
XX endothelial cell proliferation, migration, survival and differentiation
XX and lymphangiogenesis without inducing vascular permeability.
XX
SQ Sequence 109 AA;

Query Match 100.0%; Score 55; DB 20; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
DB 53 ccneeslic 61
IIIIIIIIII

RESULT 5
AA11931
ID AA11931 standard; protein: 109 AA.
XX
XX AA11931;
XX
XX 20-NOV-2000 (first entry)
XX
XX Human truncated VEGF-D.
XX
XX Truncated VEGF-D: vascular endothelial growth factor; human;
XX monoclonal antibody; VEGF receptor; VEGFR-2; VEGFR-3;
XX vascular permeability disorder; endothelial cell proliferative disorder;
XX angiogenic disorder; lymphangiogenic disorder;
XX

KW neovascularisation disorder; endothelial cell differentiation disorder;
 KW cancer; diabetic retinopathy; psoriasis; arthropathy; pulmonary oedema;
 KW detection; diagnosis; imaging; lymphatic vasculature.
 OS Homo sapiens.
 XX
 XX WO200037025-A2.
 XX
 XX 29-JUN-2000.
 XX
 XX 21-DEC-1999; 99WO-US31332.
 XX
 XX 21-DEC-1998; 98US-0113254.
 PR 17-MAY-1999; 99US-0134556.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Achen MG, Stacker SA;
 XX
 XX WPI; 2000-442498/38.
 XX
 XX Novel compositions comprising antibodies reactive to vascular
 PT endothelial growth factor-D, useful for treating e.g. angiogenesis,
 PT lymphangiogenesis and neovascularization disorders.
 XX
 XX Claim 1; Fig 1; 44pp; English.
 XX
 CC This sequence represents a 109 amino acid truncated human VEGF-D
 CC (vascular endothelial growth factor D), lacking both the N- and
 CC C-terminal regions. The invention relates to a monoclonal antibody, or
 CC fragments thereof, which is specifically reactive with the truncated
 CC human VEGF-D, and methods of preparing the antibody. The antibody of the
 CC invention interferes with the binding of VEGF-D to the VEGF receptors
 CC VEGFR-2 and VEGFR-3, but does not interfere with the binding of VEGF to
 CC these receptors and additionally is not reactive with VEGF-C. The
 CC antibody may be used to treat disorders associated with vascular
 CC permeability, endothelial cell proliferation, angiogenesis,
 CC lymphangiogenesis, neovascularisation and endothelial cell
 CC differentiation, especially cancer, diabetic retinopathy, psoriasis, and
 CC arthropathies. The antibody may also be used to treat fluid accumulation
 CC in the heart and/or lung via modulation of vascular permeability. It may
 CC additionally be used to detect VEGF-D and may be used to image lymphatic
 CC vasculature in tissue.
 XX
 XX Sequence 109 AA;
 SQ
 Query Match 100.0%; Score 55; DB 21; Length 109;
 Best Local Similarity 100.0%; Pred. No. 0.32;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCNEESLIC 9
 DB 53 cneeslic 61
 RESULT 6
 AAM53240
 ID AAM53240 standard; Protein; 325 AA.
 CC
 AC AAM53240;
 XX
 DT 03-AUG-1998 (first entry)
 XX
 XX Homo sapiens vascular endothelial growth factor D (VEGF-D).
 KW
 KW vascular endothelial growth factor; VEGF-D; angiogenesis;
 KW modification; acceleration; wound healing; tissue; organ;
 KW transplants; collateral circulation; infarction; arterial stenosis;
 KW coronary artery disease; inhibition; cancer; treatment;
 KW diabetic retinopathy; lung disorders; blood circulation;
 KW gaseous exchange; chronic obstructive airway disease;
 KW intestinal malabsorptive syndrome; biopsy; metastatic risk;

KW detection; diagnosis; congestive heart failure.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Region 126..128
 FT Region /note="potential N-linked glycosylation site"
 FT Region 156..158
 FT Region /note="potential N-linked glycosylation site"
 FT Region 258..260
 FT Region /note="potential N-linked glycosylation site"
 XX
 XX WO9807832-A1.
 XX
 XX 26-FEB-1998.
 XX
 XX 21-AUG-1997; 97WO-US14696.
 XX
 XX 01-JUL-1997; 97US-0051426.
 PR 23-AUG-1996; 96AU-0001825.
 PR 23-AUG-1996; 96US-0023751.
 PR 11-NOV-1996; 96AU-0003554.
 PR 14-NOV-1996; 96US-0031097.
 PR 05-FEB-1997; 97AU-0004954.
 PR 10-FEB-1997; 97US-0038814.
 PR 19-JUN-1997; 97AU-0007435.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 XX Achen MG, Alltalo K, Stacker SA, Wilks AF;
 XX
 XX WPI; 1998-179057/16.
 XX N-PDB; AAV20806.
 XX
 PT New isolated vascular endothelial growth factor-D - used to develop
 PT products for use in e.g. modifying angiogenesis or treating lung,
 PT heart or intestinal disorders
 XX
 XX Claim 16; Pages 57-58; 101pp; English.
 XX
 CC The sequence is that of human breast vascular endothelial growth factor
 CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
 CC in wound healing, tissue or organ transplantation, or to establish
 CC collateral circulation in tissue infarction or arterial stenosis,
 CC such as coronary artery disease, and inhibition of angiogenesis in
 CC the treatment of cancer or of diabetic retinopathy. It can also be
 CC used in the treatment of lung disorders to improve blood circulation
 CC in the lung and/or gaseous exchange between the lungs and the blood
 CC stream or to improve blood circulation to the heart and O2 gas
 CC permeability in cases of cardiac insufficiency, to improve blood
 CC flow and gaseous exchange in chronic obstructive airway disease,
 CC or to treat malabsorptive syndromes in the intestinal tract.
 CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
 CC as an indicator of future metastatic risk. Antagonists can be used
 CC for treating e.g. conditions such as congestive heart failure,
 CC involving accumulations of fluid in the lung resulting from
 CC increases in vascular permeability. The products can also be used
 CC for detection and diagnosis.
 XX
 XX Sequence 325 AA;
 SQ
 Query Match 100.0%; Score 55; DB 19; Length 325;
 Best Local Similarity 100.0%; Pred. No. 0.83;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCNEESLIC 9
 DB 116 cneeslic 124
 RESULT 7

AA97572
ID AAY97572 standard; Protein: 325 AA.
XX AAY97572;
AC AAY97572;
XX
XX
XX 05-APR-2001 (first entry)
DE Human VEGF-D protein sequence.
XX
XX Human: angiogenic protein; wound healing; vascular tissue repair;
KW peripheral arterial disease; critical limb ischaemia; coronary disease;
KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;
KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;
KW infectious disease; neurodegeneration;
KW vascular endothelial growth factor-D; VEGF-D.
XX
XX Homo sapiens.
OS
XX WO200075163-A1.
XX
XX 14-DEC-2000.
XX
XX 01-JUN-2000: 2000WO-US14925.
XX
XX 03-JUN-1999: 9905-0137796.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Hu J, Cao L;
XX WPI: 2001-071057/08.
XX N-PSDB: AAA91006.
XX
XX New nucleic acid encoding angiogenic proteins, useful e.g. for
PT promoting healing of wounds and treating peripheral arterial disease,
PT critical limb ischaemia or coronary disease -
XX
XX Claim 11: Page 226-227; 244pp; English.
XX
XX This sequence is vascular endothelial growth factor-D (VEGF-D),
XX which is an angiogenic protein of the invention. The angiogenic proteins
XX and the DNA sequences encoding them, are used to prevent, treat or
XX ameliorate disease and to detect diseases, or susceptibility, by
XX detecting mutations or the presence or amount of angiogenic protein
XX expression. Particularly they are used to stimulate wound healing,
XX growth of damaged bone and tissue, and for repair of vascular tissue,
XX especially peripheral arterial disease, critical limb ischaemia or
XX coronary disease. Antagonists of the sequences are used to inhibit
XX angiogenesis in tumours and to treat inflammation (where associated with
XX increased vascular permeability), diabetic retinopathy, rheumatoid
XX arthritis or psoriasis. Agonists are also useful for stimulating
XX (lymph)angiogenesis. The proteins are also used to identify specific
XX binding agents (potential therapeutic agents) and to raise antibodies.
XX The antibodies are useful as therapeutic (ant)agonists; for detection,
XX purification and targeting of proteins for in vivo or in vitro diagnosis
XX (including imaging) or for therapy (including when linked to e.g. a label
XX or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal
XX residual disease or haematopoietic progenitor/stem cells. It is also
XX contemplated that the sequences might be useful for treating a very wide
XX range of other disorders, e.g. autoimmune diseases; allergy; cancer;
XX infectious diseases (viral, bacterial, fungal or parasitic);
XX neurodegeneration, also as chemotactic agents or for stimulating
XX regeneration of the nervous system etc.
XX
XX Sequence 325 AA:

Query Match 100.0%; Score 55; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CNEESLIC 9
|||||||

Db 116 cneeslic 124
RESULT 8
AA49036
ID AAW49036 standard; Protein: 354 AA.
XX
XX AAW49036;
XX
XX 26-OCT-1998 (first entry)
DE Human zvegf2 growth factor.
XX
XX Human zvegf2 growth factor; fibroblast; smooth muscle cell;
KW venous stasis ulcer; diabetic ulcer; skin wound; chemotactic effect;
KW angiogenic effect; tumour; diabetic retinopathy; psoriasis; arthritis;
KW scleroderma.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23 /note= "Signal peptide"
FT 24..108 /note= "Pro-region"
FT Binding-site 109..197 /note= "Receptor binding domain"
FT Region 206..256 /note= "Cysteine-rich domain"
FT Region 257..274 /note= "Cysteine-rich domain"
FT Region 275..294 /note= "Baldiani ring motif"
FT Region 295..354 /note= "Baldiani ring motif"
FT Region /note= "Cysteine-rich domain"
XX
XX WO9824811-A2.
XX
XX 11-JUN-1998.
XX
XX 20-NOV-1997; 97WO-US20888.
XX
XX 18-SEP-1997; 97US-0933455.
XX 06-DEC-1996; 96US-0759657.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Conklin DC, Gilbert T, Hart CE, Mygaard S, Sheppard PO;
XX WPI: 1998-333256/29.
XX N-PSDB: AAV32823.
XX
XX New isolated vascular endothelial growth factor - used to develop
PT products for treating e.g. wounds, burns, myocardial infarction,
PT tumours, psoriasis, arthritis, restenosis or organ transplants
XX
XX Claim 1; Pages 53-54; 77pp; English.
XX
XX The present sequence represents a human zvegf2 growth factor encoded
XX by the zvegf2 cDNA which was isolated from a human heart cDNA library.
XX zvegf2 protein in a dimeric form acts as a mitogen for fibroblasts or
XX smooth muscle cells. zvegf2 is claimed to be useful for stimulating the
XX revascularisation of tissue or the re-endothelialisation of vascular
XX tissue. zvegf2 is particularly claimed to be useful for the treatment
XX of full-thickness skin wounds, including venous stasis ulcers and
XX diabetic ulcers. The zvegf2 protein is also claimed to be useful as an
XX additive in tissue adhesives for promoting revascularisation of the
XX healing tissue. Antagonists against zvegf2 can be used to block its
XX mitogenic, chemotactic and angiogenic effects. The antagonists may
XX therefore be useful for reducing growth of solid tumours by inhibiting
XX neovascularisation of the developing tumour or by directly blocking
XX tumour cell growth, in the treatment of diabetic retinopathy, psoriasis,
XX arthritis, and scleroderma.

```

XX Sequence      354 AA;
SQ
Query Match
Best Local Similarity 100.0%; Score 55; DB 19; Length 354;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
Db 145 ccneeslic 153

RESULT 9
AAW53241
ID AAW53241 standard; Protein; 354 AA.
AC AAW53241;
XX
XX 03-AUG-1998 (first entry)
XX
DE Homo sapiens vascular endothelial growth factor D (VEGF-D).
XX
KW vascular endothelial growth factor; VEGF-D; angiogenesis;
KW modification; acceleration; wound healing; tissue; organ;
KW transplants; collateral circulation; infarction; arterial stenosis;
KW coronary artery disease; inhibition; cancer; treatment;
KW diabetic retinopathy; lung disorders; blood circulation;
KW gaseous exchange; chronic obstructive airway disease;
KW intestinal malabsorptive syndrome; biopsy; metastatic risk;
KW detection; diagnosis; congestive heart failure.
XX
OS Homo sapiens.
XX
PN WO9807832-A1.
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14696.
XX
PR 01-JUL-1997; 97US-0051426.
PR 23-AUG-1996; 96AU-0001825.
PR 23-AUG-1996; 96US-0023751.
PR 11-NOV-1996; 96AU-0003554.
PR 14-NOV-1996; 96US-0031097.
PR 05-FEB-1997; 97AU-0004954.
PR 10-FEB-1997; 97US-0038814.
PR 19-JUN-1997; 97AU-0007435.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD.
XX
PI Aachen MG, Altalo K, Stacker SA, Wilks AF;
XX
XX WPI: 1998-179057/16.
DR N-PSDB; AAV20807.
XX
XX New Isolated vascular endothelial growth factor-D - used to develop
PT products for use in e.g. modifying angiogenesis or treating lung,
PT heart or intestinal disorders
XX
XX Claim 16; Pages 60-61; 101pp; English.
XX
XX The sequence is that of human lung vascular endothelial growth factor
CC D (VEGF-D). VEGF-D can be used for e.g. acceleration of angiogenesis
CC in wound healing, tissue or organ transplantation, or to establish
CC collateral circulation in tissue infarction or arterial stenosis,
CC such as coronary artery disease, and inhibition of angiogenesis in
CC the treatment of cancer or of diabetic retinopathy. It can also be
CC used in the treatment of lung disorders to improve blood circulation
CC in the lung and/or gaseous exchange between the lungs and the blood
CC stream or to improve blood circulation to the heart and O2 gas
CC permeability in cases of cardiac insufficiency, to improve blood

```

```

CC flow and gaseous exchange in chronic obstructive airway disease,
CC or to treat malabsorptive syndromes in the intestinal tract.
CC Quantitation of VEGF-D in cancer biopsy specimens may be useful
CC as an indicator of future metastatic risk. Antagonists can be used
CC for treating e.g. conditions such as congestive heart failure,
CC involving accumulations of fluid in the lung resulting from
CC increases in vascular permeability. The products can also be used
CC for detection and diagnosis.
XX
SQ Sequence      354 AA;

Query Match
Best Local Similarity 100.0%; Score 55; DB 19; Length 354;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
Db 145 ccneeslic 153

RESULT 10
AAW44293
ID AAW44293 standard; Protein; 354 AA.
AC AAW44293;
XX
XX 22-JUN-1998 (first entry)
XX
DE Human vascular endothelial growth factor D.
XX
KW Human; vascular endothelial growth factor D; VEGF-D; gene therapy;
KW inflammation; oedema.
XX
OS Homo sapiens.
XX
PN WO9802543-A1.
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-JP02456.
XX
PR 15-JUL-1996; 96JP-0185216.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Hirata Y, Nezu J;
XX
XX WPI: 1998-110591/10.
DR N-PSDB; AAV15156.
XX
XX VEGF-D protein encoded by DNA - useful for, e.g. gene therapy and
PT treating oedema
XX
XX Claim 1; Page 18-20; 52pp; Japanese.
XX
XX The present sequence represents human vascular endothelial growth factor
CC D (VEGF-D). The VEGF-D protein, compounds and antibodies, which can bind
CC the protein, may be useful in, e.g. gene therapy and in treatment of
CC inflammation and oedema. Vectors, containing the VEGF-D DNA, and VEGF-D
CC DNA sequences may be used for screening for the compounds which bind to
CC the VEGF-D protein.
XX
SQ Sequence      354 AA;

Query Match
Best Local Similarity 100.0%; Score 55; DB 19; Length 354;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCNEESLIC 9
Db 145 ccneeslic 153

```


KM	Human; receptor tyrosine kinase: RTK; Flt4; fms-like tyrosine kinase 4;
KV	VEGFR-3; vascular endothelial growth factor receptor-3; chromosome 5q35;
KW	cytostatic; tumour imaging; anti-tumour therapy; treatment; diagnosis;
KX	neoplastic disease; lymphoma; carcinoma; breast; squamous cell; melanoma;
KY	sarcoma; malignancy; VEGF-D; vascular endothelial growth factor D.
OS	Homo sapiens.
PN	WO20021560-A1.
PD	20-APR-2000.
PE	08-OCT-1999; 99WO-US23525.
PR	09-OCT-1998; 98US-0169079.
PA	(LUDMW-) LUDMWG INST CANCER RES.
XX	(UYHE-) UNIV HELSINKI LICENSING LTD OY.
PI	Allitalo K, Kaipainen A, Vaitola R, Jussila L;
DR	WPI, 2000-317850/27.
PS	Example 15-17; Page 142-143; 148pp; English.
XX	The patent discloses a method to treat neoplastic disease characterised
CC	by expression of fms-like tyrosine kinase 4 (Flt4) receptor (also
CC	referred as vascular endothelial growth factor receptor-3, VEGFR-3) in
CC	endothelial cells of blood vessels adjacent to malignant neoplasm. The
CC	method involves administering a compound that inhibits binding of a
CC	ligand to Flt4 thereby inhibiting Flt4 mediated proliferation of vascular
CC	endothelial cells. The compound is useful for treating neoplastic disease
CC	such as breast carcinomas, squamous cell carcinomas, lymphomas, melanomas
CC	and sarcomas. Flt4 receptor tyrosine kinase binding compounds can be used
CC	for manufacturing medicament useful for diagnostic screening, imaging and
CC	treatment of malignancies characterised by Flt4-expressing blood cells.
CC	The Flt4 gene maps to chromosomal region 5q35 and is expressed as 5.8 kb
CC	and 4.5 kb mRNAs which differ in their 3' sequences and are
CC	differentially expressed in HEL and DAMI cell lines. Flt4
CC	belongs to a subfamily of class III receptor tyrosine kinases (RTKs).
CC	It is used as a target for tumour imaging and anti-tumour therapy.
CC	The present sequence is a human prepro-vascular endothelial growth
CC	factor D (VEGF-D), a specific example of Flt4 binding compound. A
CC	recombinantly matured VEGF-D lacking residues 1-92 and 202-354
CC	retains the ability to activate VEGFR-2 and VEGFR-3 receptors and
CC	associate as non-covalently linked dimers.
XX	
SQ	Sequence 354 AA:
OY	1 CCNESSLIC 9
DB	145 CCNESSLIC 153
ID	AAY70983 standard; Protein: 354 AA.
AC	AAY70983:
XX	
DT	09-JUG-2000 (first entry)
DE	Human vascular endothelial growth factor (VEGF)-D protein.
XX	

KV	Vascular endothelial growth factor-D; VEGF; human; re-endothelialisation;
KW	vascular endothelial growth factor receptor; VEGFR; vascular trauma;
KM	blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;
KX	restenosis; stenosis; percutaneous transluminal coronary angioplasty.
XX	
OS	Homo sapiens.
FH	Key
FT	Peptide
FT	1..21
FT	/label= Signal_peptide
FT	/note= "Cleavage results in partially-processed VEGF-D protein"
FT	22..92
FT	/label= Amino_terminal_peptide
FT	/note= "Cleavage results in fully-processed mature VEGF-D protein"
FT	93..201
FT	/label= Recombinantly_matured_VEGF_D_protein
FT	/note= "Processed vascular epithelial growth factor-D"
FT	202..354
FT	/label= Carboxy_terminal_peptide
FT	/note= "Cleavage results in partially-processed VEGF-D protein"
FN	
XX	WO200024412-A2.
PD	04-MAY-2000.
XX	
PF	26-OCT-1999; 99WO-US24054.
XX	
PR	26-OCT-1998; 98US-0105587.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
PA	(UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX	(YLAH/) YLA-HERTTUALA S.
PI	Yla-herttuala S., Haltalo K., Hiltunen MO., Jeltsch MM., Achen MG;
XX	
DR	WPI: 2000-350584/30.
DR	N-PSDB; AAD00340.
XX	
PT	Preventing stenosis and restenosis in mammals using vascular
XX	endothelial growth factor proteins or the nucleic acids encoding them -
PS	Disclosure; Page 53-55; 61pp; English.
XX	
CC	The present amino acid sequence is the complete human prepro-vascular
CC	endothelial growth factor (VEGF)-D. VEGF-D has the ability to stimulate
CC	re-endothelialisation of an injured blood vessel, without significant
CC	stimulation of smooth muscle cell proliferation. It can bind to and/or
CC	stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or
CC	VEGFR-3 phosphorylation in cells that express such receptors. An
CC	anti-restenosis agent comprising either a VEGF-D gene or protein is
CC	used in a method to reduce or prevent restenosis and stenosis of a blood
CC	vessel following vascular trauma e.g., cardiovascular surgery and
CC	percutaneous transluminal coronary angioplasty.
XX	
SQ	Sequence 354 AA:
OY	Query Match 100.0%; Score 55; DB 21; Length 354;
	Best Local Similarity 100.0%; Pred. No. 0.9;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 CCNEESLIC 9
	145 ccneeslic 153
RESULT 15	
AAU08441	
ID	AAU08441 standard; Protein: 354 AA.
XX	

AC AAU08441:

XX 21-NOV-2001 (first entry)

XX Polypeptide for human VEGF-D.

DE Human: vascular endothelial growth factor; VEGF-D; vasculogenesis;
 XX angiogenesis; blood vessel; cancer; proliferative retinopathy;
 KW psoriasis; age-related macular degeneration; rheumatoid arthritis;
 KW cardiovascular.

XX Homo sapiens.

XX WO200162942-A2.

XX 30-AUG-2001.

XX 26-FEB-2001: 2001WO-US06113.

XX 25-FEB-2000: 2000US-0185205.

XX 18-MAY-2000: 2000US-0205331.

XX (LUDWIG-) LUDWIG INST CANCER RES.

XX (LICHT) LICHTA OY.

XX Altalo K, Jeltsch KM;

XX WPI: 2001-536640/59.

XX N-PSDB; AAS12878.

XX Polypeptides that bind cellular receptors for vascular endothelial

XX growth factors, polynucleotides encoding them -

XX Disclosure: Page 227-228; 261pp; English.

XX The present invention relates to polypeptides that bind cellular
 CC receptors for vascular endothelial growth factors (VEGFs), the
 CC polynucleotides encoding them, and their use for identifying agents that
 CC modulate interactions between VEGFs and their receptors. VEGFs and their
 CC receptors play an important role in vasculogenesis, the development of
 CC the embryonic vasculature from early differentiating endothelial cells
 CC and angiogenesis, the process of forming new blood vessels from
 CC pre-existing ones. Modulators of interactions between VEGF and its
 CC receptors may be used to treat dysfunction of the endothelial cell
 CC regulatory system. Such disorders include cancers, abnormal angiogenesis,
 CC proliferative retinopathies, age-related macular degeneration, rheumatoid
 CC arthritis and psoriasis. The polypeptides of the invention exhibit unique
 CC receptor binding profiles compared to known naturally occurring VEGFs.
 CC The present sequence represents the polypeptide for human VEGF-D.

XX Sequence 354 AA;

Query Match

100.0%; Score 55; DB 22; Length 354;

Best Local Similarity 100.0%; Pred. No. 0.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9

DB 145 ccneeslic 153

Search completed: June 17, 2002, 16:02:13
 Job time: 422 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:37 ; Search time 75.98 Seconds
(without alignments)
2.893 Million cell updates/sec

Title: US-09-761-636A-9
Perfect score: 55
Sequence: 1 CCNESLIC 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	325	4	US-08-915-795-3
2	55	100.0	354	4	US-08-915-795-5
3	45	81.8	120	6	5194596-9
4	45	81.8	120	6	5219739-9
5	45	81.8	164	6	5194596-17
6	45	81.8	164	6	5219739-17
7	45	81.8	164	6	5219739-18
8	45	81.8	190	6	5332671-3
9	45	81.8	321	4	US-08-915-795-9
10	45	81.8	358	4	US-08-915-795-8
11	42	76.4	146	1	US-08-586-039B-33
12	42	76.4	189	1	US-08-469-427A-15
13	42	76.4	190	2	US-08-569-063C-20
14	42	76.4	190	4	US-08-586-039B-31
15	42	76.4	214	4	US-08-586-039B-35
16	41	74.5	12	3	US-08-742-243-43
17	41	74.5	12	3	US-08-742-243-44
18	41	74.5	109	3	US-08-691-794-3
19	41	74.5	110	4	US-09-392-932-11
20	41	74.5	121	6	5194596-19
21	41	74.5	121	6	5219739-20
22	41	74.5	145	3	US-08-784-551C-2
23	41	74.5	145	4	US-09-392-932-2
24	41	74.5	147	3	US-08-807-992B-1
25	41	74.5	147	4	US-09-392-932-1
26	41	74.5	165	6	5194596-18
27	41	74.5	165	6	5219739-19

28	41	74.5	191	3	US-08-567-200A-2	Sequence 2, Appl1
29	41	74.5	191	3	US-08-807-992B-2	Sequence 2, Appl1
30	41	74.5	191	3	US-08-691-794-2	Sequence 2, Appl1
31	41	74.5	191	4	US-08-795-430-56	Sequence 56, Appl1
32	41	74.5	191	4	US-09-392-932-3	Sequence 3, Appl1
33	41	74.5	191	6	5332671-4	Patent No. 5332671
34	41	74.5	214	6	5240848-11	Patent No. 5240848
35	41	74.5	215	3	US-08-807-992B-3	Sequence 3, Appl1
36	41	74.5	215	4	US-08-586-039B-49	Sequence 49, Appl1
37	41	74.5	215	6	5219739-22	Patent No. 5219739
38	41	74.5	215	6	5240848-7	Patent No. 5240848
39	41	74.5	231	5	PCR-US96-09001-10	Sequence 10, Appl1
40	41	74.5	232	2	US-08-999-811-7	Sequence 7, Appl1
41	41	74.5	232	2	US-08-824-996-9	Sequence 9, Appl1
42	41	74.5	232	3	US-08-807-992B-4	Sequence 4, Appl1
43	41	74.5	232	3	US-09-042-105-7	Sequence 7, Appl1
44	39	70.9	350	2	US-08-999-811-4	Sequence 4, Appl1
45	39	70.9	350	2	US-08-824-996-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Karl ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKewen, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Breast
US-08-915-795-3
Query Match 100.0%; Score 55; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
|||||||

Db 116 CCNESLIC 124

RESULT 2
US-08-915-795-5
; Sequence 5, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
; US-08-915-795-5

Query Match 100.0%; Score 55; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
|||||||

Db 145 CCNESLIC 153

RESULT 3
5194596-9
; Patent No. 5194596
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN
; C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/450,883
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO: 9;
; LENGTH: 120
5194596-9

Query Match 81.8%; Score 45; DB 6; Length 120;
Best Local Similarity 77.8%; Pred. No. 2.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
|||||||

Db 59 CCNESLIC 67

RESULT 4
5219739-9
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGF120 AND
; BVGF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGF120 AND BVGF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO: 9;
; LENGTH: 120
5219739-9

Query Match 81.8%; Score 45; DB 6; Length 120;
Best Local Similarity 77.8%; Pred. No. 2.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
|||||||

Db 59 CCNESLIC 67

RESULT 5
5194596-17
; Patent No. 5194596
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES, JOHN
; C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/450,883
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO: 17;
; LENGTH: 164
5194596-17

Query Match 81.8%; Score 45; DB 6; Length 164;
Best Local Similarity 77.8%; Pred. No. 3.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
111:1111
Db 59 CCNDESLC 67

RESULT 6
5219739-17
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; HVGEF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVGEF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:17:
; LENGTH: 164
5219739-17

Query Match 81.8%; Score 45; DB 6; Length 164;
Best Local Similarity 77.8%; Pred. No. 3.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCNEESLIC 9
111:1111
Db 59 CCNDESLC 67

RESULT 7
5219739-18
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.; ABRAHAM, JUDITH A.; FIDDES,
; JOHN C.; MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; HVGEF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVGEF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450,883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:18:
; LENGTH: 164
5219739-18

Query Match 81.8%; Score 45; DB 6; Length 164;
Best Local Similarity 77.8%; Pred. No. 3.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCNEESLIC 9
111:1111
Db 59 CCNDESLC 67

RESULT 8
5332671-3
; Patent No. 5332671
; APPLICANT: FERRARA, NAPOLITANO; LEUNG, DAVID W.H.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR AND DNA ENCODING SAME

NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/389,722
FILING DATE: 04-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 369,424
FILING DATE: 21-JUN-1989
APPLICATION NUMBER: 351,117
FILING DATE: 12-MAY-1989
SEQ ID NO:3:
LENGTH: 190
5332671-3

Query Match 81.8%; Score 45; DB 6; Length 190;
Best Local Similarity 77.8%; Pred. No. 4.2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCNEESLIC 9
111:1111
Db 85 CCNDESLC 93

RESULT 9
US-08-915-795-9
; Sequence 9, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Karl ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; TISSUE TYPE: Mouse Lung
; US-08-915-795-9

Query Match 81.8%; Score 45; DB 4; Length 321;
Best Local Similarity 66.7%; Pred. No. 6.7;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNESLIC 9
Db 145 CCNEGVMC 153

RESULT 10

US-08-915-795-8

Sequence 8, Application US/08915795

Patent No. 623713

GENERAL INFORMATION:

APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS

APPLICANT: Steven A. STACKER

APPLICANT: Karl ALITALO

TITLE OF INVENTION: GROWTH FACTOR

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.

STREET: 1200 G Street, NW, Suite 700

CITY: Washington

STATE: DC

COUNTRY: United States of America

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915.795

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 1064/42983

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

TELEX: N/A

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 358 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

TISSUE TYPE: Mouse Lung

US-08-915-795-8

QY 1 CCNESLIC 9
Db 150 CCNEGVMC 158

Query Match 81.8%; Score 45; DB 4; Length 358;

Best Local Similarity 66.7%; Pred. No. 7.4;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 11

US-08-586-039B-33

Sequence 33, Application US/08586039B

Patent No. 6140073

GENERAL INFORMATION:

APPLICANT: Bayne, Marvin L.

APPLICANT: Thomas, Jr., Kenneth A.

TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 E. Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Microsoft Word 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/586.039B

FILING DATE: 16-JAN-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/124,259

FILING DATE: 20-SEP-1993

APPLICATION NUMBER: 07/676,436

FILING DATE: 28-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hand, J. Mark

REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 18361DA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3905

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 146 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-586-039B-33

QY 1 CCNESLIC 9
Db 85 CCNDEALEC 93

Query Match 76.4%; Score 42; DB 4; Length 146;

Best Local Similarity 66.7%; Pred. No. 9.2;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 12

US-08-469-427A-15

Sequence 15, Application US/08469427A

Patent No. 5607918

GENERAL INFORMATION:

APPLICANT: Eriksson, Ulf

APPLICANT: Olofsson, Birgitta

APPLICANT: Alitalo, Kari

TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469.427A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,651
FILING DATE: 01-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 41979CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-427A-15

Query Match 76.4% Score 42; DB 1; Length 189;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
111:111
DB 84 CCNDEALEC 92

RESULT 13
US-08-569-063C-20
Sequence 20, Application US/08569063C
Patent No. 5928939
GENERAL INFORMATION:
APPLICANT: ERIKSSON, Ulf
APPLICANT: OLOFSSON, Birgitta
APPLICANT: ALITALO, Kari
APPLICANT: PAJUSOLA, Kari
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
TITLE OF INVENTION: DNA CODING THEREFOR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKown, Edwards & Lenahan, P.L.L.C.
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,063C
FILING DATE: 06-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/469,427
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/397,651
FILING DATE: 01-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/41979CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-063C-20

Query Match 76.4% Score 42; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
111:111
DB 85 CCNDEALEC 93

RESULT 14
US-08-586-039B-31
Sequence 31, Application US/08586039B
Patent No. 614073
GENERAL INFORMATION:
APPLICANT: Bayne, Marvin L.
APPLICANT: Thomas Jr., Kenneth A.
TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C
TITLE OF INVENTION: SUBUNIT
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 E. Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word 6

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/586,039B
FILING DATE: 16-JAN-1996

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/124,259
FILING DATE: 20-SEP-1993
APPLICATION NUMBER: 07/676,436
FILING DATE: 28-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Hand, J. Mark
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18361DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4720
TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-586-039B-31

Query Match 76.4% Score 42; DB 4; Length 190;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCNEESLIC 9
111:111
DB 85 CCNDEALEC 93

RESULT 15

US-08-586-039B-35
; Sequence 35, Application US/08586039B
; Patent No. 6140073
; GENERAL INFORMATION:
; APPLICANT: Bayne, Marvin L.
; APPLICANT: Thomas Jr., Kenneth A.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL CELL GROWTH FACTOR C
; TITLE OF INVENTION: SUBUNIT
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 E. Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,039B
; FILING DATE: 16-JAN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/124,259
; FILING DATE: 20-SEP-1993
; APPLICATION NUMBER: 07/676,436
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 18361DA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-586-039B-35

Query Match 76.4%; Score 42; DB 4; Length 214;

Best Local Similarity 66.7%; Pred. No. 13;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CNEESLIC 9
|||:|:|
Db 85 CCNDEALC 93

Search completed: June 17, 2002, 15:56:37
Job time: 86 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:23 ; Search time 95.43 seconds
(without alignments)
9.062 Million cell updates/sec

Title: US-09-761-636A-10

Perfect score: 50

Sequence: 1 CSVPLTSVC 9

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	86.0	1487	2	phospholipase-A(2)
2	41	82.0	349	2	probable membrane
3	41	82.0	364	2	hypothetical prote
4	41	82.0	376	2	hypothetical 38.5k
5	41	82.0	376	2	hypothetical prote
6	39	78.0	432	2	probable transamin
7	38	76.0	22	2	uvrc protein - pse
8	36	72.0	591	2	calcium-stimulated
9	36	72.0	2180	2	hypothetical prote
10	35	70.0	91	2	transcription regu
11	35	70.0	124	2	napin - Swedish tu
12	35	70.0	127	2	allergen Sin a I -
13	35	70.0	133	1	napin I precursor
14	35	70.0	145	2	allergen Sin a I
15	35	70.0	145	2	allergen Sin a I
16	35	70.0	145	2	allergen Sin a I
17	35	70.0	145	2	allergen Sin a I
18	35	70.0	152	2	hypothetical prote
19	35	70.0	155	2	napin B9 precursor
20	35	70.0	155	2	napin 2 precursor
21	35	70.0	178	1	napin B - rape
22	35	70.0	178	2	napin precursor (n
23	35	70.0	178	2	2S storage protein
24	35	70.0	178	2	2S storage protein
25	35	70.0	178	2	2S storage protein
26	35	70.0	180	2	napin (clone BngNA
27	35	70.0	180	2	napin (clones BNM
28	35	70.0	621	1	monophenol monooxy
29	34	68.0	92	2	major allergen lar

30	34	68.0	141	2	T12116
31	34	68.0	178	2	S15382
32	34	68.0	192	2	B96383
33	34	68.0	265	2	T11729
34	34	68.0	270	2	T11687
35	34	68.0	273	2	T11702
36	34	68.0	277	2	T11697
37	34	68.0	278	2	T11731
38	34	68.0	376	2	S70841
39	34	68.0	377	2	S77634
40	34	68.0	386	2	A57066
41	34	68.0	418	2	E71716
42	34	68.0	565	2	T08794
43	34	68.0	659	1	JC4365
44	34	68.0	661	1	UN0870
45	34	68.0	669	2	T17122

ALIGNMENTS

RESULT 1
S48719
phospholipase-A(2) receptor protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C:Accession: S48719
R:Higashino, K.; Ishizaki, J.; Kishino, J.; Ohara, O.; Arita, H.
Eur. J. Biochem. 225: 375-382, 1994
A:Title: Structural comparison of phospholipase-A(2)-binding regions in phospholip
A:Reference number: S48719; MUID:95010128
A:Accession: S48719
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 11467 <HIG>
A:Cross-references: GB:030779; NID:91375042; PIDN:BAA06443.1; PID:9691754
C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type
F:181-222/Domain: fibronectin type II repeat homology <ZFI>
F:380-503/Domain: C-type lectin homology <LCH>

Query Match 86.0%; Score 43; DB 2; Length 1487;
Best Local Similarity 77.8%; Pred. No. 8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CSVPLTSVC 9
Db 929 CSVPLTSVC 937
RESULT 2
AD0990
probable membrane protein STY4229 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0990
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Chum
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fe
S.; Mout, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Steven
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0990
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08048.1; PID:916505028; GSPDB:GN00176
C:Genetics:
A:Gene: STY4229
C:Superfamily: conserved hypothetical protein H10338

Query Match 82.0%; Score 41; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSVC 9
 DB 314 SVPLTSVC 321

RESULT 3

hypothetical protein yhhT [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: C86015
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grobleck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C86015
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <STO>
 A:Cross-references: GB:AE005174; NID:91251813; PIDN:AAG58583.1; GSPDB:GN00145; UWGP:Z4B
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: yhhT
 C:Superfamily: conserved hypothetical protein HI0338

Query Match 82.0%; Score 41; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSVC 9
 DB 329 SVPLTSVC 336

RESULT 4

hypothetical 38.5K protein (fsy-nika intergenic region) - Escherichia coli
 N:Alternate names: hypothetical protein o376
 C:Species: Escherichia coli
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 29-Sep-1999
 C:Accession: S47693; E65144
 R:Plunkett, G.
 submitted to the EMBL Data Library, March 1994
 A:Reference number: S47666
 A:Accession: S47693
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <PLU>
 A:Cross-references: EMBL:U00039; NID:9466582; PIDN:AB18449.1; PID:912460
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: E65144
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-376 <BLAT>
 A:Cross-references: GB:AE000423; GB:U00096; NID:91789880; PIDN:AAC76499.1; PID:91789885;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: yhhT
 A:Start codon: GNG
 C:Superfamily: conserved hypothetical protein HI0338

Query Match 82.0%; Score 41; DB 2; Length 376;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSVC 9
 DB 341 SVPLTSVC 348

RESULT 5

hypothetical protein ECs4323 [imported] - Escherichia coli (strain O157:H7, substrain
 C91169
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: C91169
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: C91169
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA37746.1; PID:913363797; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs4323
 C:Superfamily: conserved hypothetical protein HI0338

Query Match 82.0%; Score 41; DB 2; Length 376;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSVC 9
 DB 341 SVPLTSVC 348

RESULT 6

probable transaminase (EC 2.6.1.-) MJ1479 [similarity] - Methanococcus jannaschii
 F64484
 N:Alternate names: alanine aminotransferase 2
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jun-2000
 C:Accession: F64484
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Bl
 ; Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek
 son, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woe
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus janna
 A:Reference number: A64300; MUID:96337999
 A:Accession: F64484
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-432 <BLU>
 A:Cross-references: GB:U67586; GB:L77117; NID:91592111; PID:91592118; TIGR:MJ1479;
 A:Map position: FOR1449710-1451008
 C:Genetics:
 A:Superfamily: aspartate transaminase
 C:Keywords: aminotransferase

Query Match 78.0%; Score 39; DB 2; Length 432;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPLTSVC 9
 DB 389 CSVPLTSVC 397

RESULT 7
B5538
uvrC protein - Pseudomonas syringae pv. syringae (strain B728a) (fragment)
C:Species: Pseudomonas syringae pv. syringae
C:Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 08-Sep-1995
C:Accession: B5538
R:Alch, J.J.; Kinscherf, T.G.; Kitten, T.; Willis, D.K.
J. Bacteriol. 176, 7468-7475, 1994
A:Title: Genetic evidence that the gacA gene encodes the cognate response regulator for
A:Reference number: A5538; MUID:95095914
A:Accession: B5538
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <RUC>
A:Cross-references: GB:009767
C:Genetics:
A:Gene: uvrC

Query Match 76.0%; Score 36; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.4;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 CSVPITSV 9
||| :||
Db 13 CSVPITSV 21

RESULT 8
S54788
calcium-stimulated protein kinase - Chlamydomonas eugametos
C:Species: Chlamydomonas eugametos
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1995
C:Accession: S54788
R:Siderius, M.H.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning a CDPK from Chlamydomonas eugametos.
A:Reference number: S54788
A:Accession: S54788
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-591 <SID>
A:Cross-references: EMBL:249233; NID:9806541; PID:CA89202.1; PID:9806542
C:Keywords: calcium-dependent protein kinase; calmodulin repeat homology; protein kin
C:Keywords: ATP; calcium binding; EF hand; serine/threonine-specific protein kinase; tan
F:150-410/Domain: protein kinase homology <KIN>
F:158-166/Region: protein kinase ATP-binding motif
F:453-485/Domain: calmodulin repeat homology <EF1>
F:489-521/Domain: calmodulin repeat homology <EF2>
F:525-557/Domain: calmodulin repeat homology <EF3>
F:559-591/Domain: calmodulin repeat homology <EF4>

Query Match 72.0%; Score 36; DB 2; Length 591;
Best Local Similarity 55.6%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSVPITSV 9
| :||| |
Db 88 CKIPITRAC 96

RESULT 9
T29764
hypothetical protein T21E3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T29764
R:Du, Z.; Le, T.T.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid T21E3.
A:Reference number: Z20681

A:Accession: T29764
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2180 <DNZ>
A:Cross-references: EMBL:AF003133; PID:AA54138.1; GSPDB:GN00019; CESP:T21E3.3
C:Genetics:
A:Gene: CESP:T21E3.3
A:Map position: 1
A:Interons: 400/2; 1608/3; 1644/1; 1674/3; 1686/2; 1731/2; 2073/1; 2098/2; 2146/3
C:Superfamily: LDL receptor ligand-binding repeat homology <LDL1>
F:15-49/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:51-85/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:90-132/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:149-185/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:190-225/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:228-272/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:285-316/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:323-359/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:369-405/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:816-856/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:861-903/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:908-943/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F:948-987/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F:993-1027/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F:1031-1063/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F:1126-1165/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F:1172-1209/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F:1213-1246/Domain: LDL receptor ligand-binding repeat homology <LDL18>

Query Match 72.0%; Score 36; DB 2; Length 2180;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VPIITSV 9
:|||||
Db 922 IPIITSV 928

RESULT 10
D97130
transcription regulator (phage-related) (Xre family) [imported] - Clostridium acet
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: D97130
R:Nolling, J.; Bennett, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacteri
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97130
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <KUR>
A:Cross-references: GB:AE001437; PID:AAK79831.1; PID:q15024845; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1867

Query Match 70.0%; Score 35; DB 2; Length 91;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITSV 9
|| :||| |
Db 54 CSITITSV 62

RESULT 11
A37931
napin - Swedish turnip (fragments)

C:Species: Brassica napus var. rapifera (Swedish turnip, rutabaga)
C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995
C:Accession: A37931
R:Syndsen, I.; Nicolova, D.; Goshiev, I.; Genov, N.
Carlsberg Res. Commun. 54, 231-239, 1989
A:Title: Isolation and characterization of a trypsin inhibitor from the seeds of kohlrabi
A:Reference number: A37931; MUID:91282906
A:Accession: A37931
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-124 <SVE>
C:Superfamily: wheat alpha-amylase inhibitor

Query Match 70.0%; Score 35; DB 2; Length 124;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPITVVC 9
DB 106 CNIPQVSVVC 114

RESULT 12
S65447
allergen Sin a I - white mustard
N:Alternate names: allergen Sin a I large chain
C:Species: Sinapis alba (white mustard)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
R:Gonzalez de la Pena, M.A.; Monsalve, R.I.; Batanero, E.; Villalba, M.; Rodriguez, R.
Eur. J. Biochem. 237, 827-832, 1996
A:Title: Expression in Escherichia coli of Sin a I, the major allergen from mustard.
A:Reference number: S65447; MUID:96235251
A:Accession: S65447
A:Molecule type: protein
A:Residues: 1-127 <GON>
A:Experimental source: seed
R:Menendez-Arias, L.; Moneo, I.; Dominguez, J.; Rodriguez, R.
Eur. J. Biochem. 177, 159-166, 1988
A:Title: Primary structure of the major allergen of yellow mustard (Sinapis alba L.) seed
A:Reference number: S01791; MUID:89030681
A:Accession: S01792
A:Molecule type: protein
A:Residues: 40-127 <MEN>
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein

Query Match 70.0%; Score 35; DB 2; Length 127;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPITVVC 9
DB 109 CNIPQVSVVC 117

RESULT 13
NMRP1
napin 1 precursor - rape (fragment)
N:Alternate names: 1.7S seed storage protein
C:Species: Brassica napus (rape)
C:Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 18-Jun-1999
R:Crouch, M.L.; Tenbarger, K.M.; Simon, A.E.; Ferl, R.
J. Mol. Appl. Genet. 2, 273-285, 1983
A:Title: cDNA clones for Brassica napus seed storage proteins: evidence from nucleotide
A:Reference number: A92836; MUID:84113267
A:Accession: A01330
A:Molecule type: mRNA
A:Residues: 1-133 <CRO>
A:Cross-references: GB:K01544; NID:g167174; PIDN:AAA33005.1; PID:g167175

A:Experimental source: cv. Tower
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein
F:1-30/Product: napin 1 small chain (fragment) #status predicted <SCH>
F:50-130/Product: napin 1 large chain #status predicted <LCH>

Query Match 70.0%; Score 35; DB 1; Length 133;
Best Local Similarity 55.6%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPITVVC 9
DB 114 CNIPQVSVVC 122

RESULT 14
S65478
allergen Sin a I (clone SIN5) - white mustard (fragment)
C:Species: Sinapis alba (white mustard)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 16-Jul-1999
R:Gonzalez de la Pena, M.A.; Monsalve, R.I.; Batanero, E.; Villalba, M.; Rodriguez, R.
Eur. J. Biochem. 237, 827-832, 1996
A:Title: Expression in Escherichia coli of Sin a I, the major allergen from mustard
A:Reference number: S65478; MUID:96235251
A:Accession: S65478
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-145 <GON>
A:Cross-references: EMBL:X91798; NID:g1009441; PIDN:CAA62908.1; PID:g1009442
A:Experimental source: seed
C:Genetics: sin5
A:Gene: sin5
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein

Query Match 70.0%; Score 35; DB 2; Length 145;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPITVVC 9
DB 127 CNIPQVSVVC 135

RESULT 15
S65481
allergen Sin a I (clone SIN3) - white mustard (fragment)
C:Species: Sinapis alba (white mustard)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 16-Jul-1999
R:Gonzalez de la Pena, M.A.; Monsalve, R.I.; Batanero, E.; Villalba, M.; Rodriguez, R.
Eur. J. Biochem. 237, 827-832, 1996
A:Title: Expression in Escherichia coli of Sin a I, the major allergen from mustard
A:Reference number: S65447; MUID:96235251
A:Accession: S65481
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-145 <GON>
A:Cross-references: EMBL:X91801; NID:g1009437; PIDN:CAA62911.1; PID:g1009438
A:Experimental source: seed
C:Genetics: sin3
A:Gene: sin3
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein

Query Match 70.0%; Score 35; DB 2; Length 145;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Mon Jun 17 16:10:55 2002

us-09-761-636a-10.open.rpr

Page 5

Oy 1 CSVPLTSVC 9
1::1 111
Db 127 CNIPQV/SVC 135

Search completed: June 17, 2002, 15:58:24
Job time: 193 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:28 ; Search time 44.84 Seconds
(Without alignments)
7.772 Million cell updates/sec

Title: US-09-761-636A-10

Perfect score: 50

Sequence: 1 CSVPLTSVC 9

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	82.0	349	1 YHHT_ECOLI	P37622 escherichia
2	38	76.0	150	1 UVRC_PSESY	O52377 pseudomonas
3	35	70.0	127	1 ALLL_STNAL	P15322 sinapis alb
4	35	70.0	133	1 2SS1_BRANA	P01091 brassica na
5	35	70.0	178	1 2SS2_BRANA	P01090 brassica na
6	35	70.0	180	1 2SS3_BRANA	P17333 brassica na
7	35	70.0	620	1 TYRO_NEUCR	P00440 neurospora
8	34	68.0	129	1 ALL1_BRAJU	P80207 brassica ju
9	34	68.0	178	1 2SSB_BRANA	P27740 brassica na
10	34	68.0	377	1 EPA2_RALSO	O45407 talstonia s
11	34	68.0	386	1 P1ZL_HUMAN	P43119 homo sapien
12	34	68.0	574	1 CIRS_HUMAN	O43734 homo sapien
13	34	68.0	659	1 SYR_HUMAN	P54136 homo sapien
14	34	68.0	661	1 SYR_HUMAN	P37880
15	33	66.0	100	1 GARC_RITCN	O92374 tricketulus
16	33	66.0	311	1 HEM3_NEIMA	O92374 tricketulus
17	33	66.0	311	1 HEM3_NEIMA	O92374 tricketulus
18	33	66.0	572	1 NUSM_LOCOMI	O92374 tricketulus
19	33	66.0	1458	1 PAZL_RABIT	O92374 tricketulus
20	32	64.0	178	1 YGV5_YEAST	O92374 tricketulus
21	32	64.0	216	1 NKGD_HUMAN	O92374 tricketulus
22	32	64.0	308	1 SAL_STILAS	O92374 tricketulus
23	32	64.0	326	1 VEGD_RAT	O92374 tricketulus
24	32	64.0	354	1 VEGD_HUMAN	O92374 tricketulus
25	32	64.0	358	1 VEGD_MOUSE	O92374 tricketulus
26	32	64.0	378	1 ADHQ_RABIT	O92374 tricketulus
27	32	64.0	391	1 ADHQ_HUMAN	O92374 tricketulus
28	32	64.0	403	1 G3PA_MAIZE	O92374 tricketulus
29	32	64.0	427	1 G6WT_BOVIN	O92374 tricketulus
30	32	64.0	451	1 G3PB_PEA	O92374 tricketulus
31	32	64.0	598	1 NUSM_BRALA	O92374 tricketulus
32	32	64.0	598	1 Y427_HUMAN	O92374 tricketulus
33	32	64.0	599	1 NUSM_BRAFL	O92374 tricketulus

34	32	64.0	660	1 VNC5_PAVPN	P18547 porcine par
35	32	64.0	662	1 VNC5_PAVK	P52502 porcine par
36	32	64.0	1115	1 IREL_YEAST	P32361 saccharomyc
37	32	64.0	1463	1 PAZL_BOVIN	P49259 bos taurus
38	31	62.0	130	1 ITRY_SINAR	P38057 sinapis arv
39	31	62.0	144	1 NOGG_RAT	O62809 ratius norv
40	31	62.0	149	1 DISD_DICDI	P02888 dictyostell
41	31	62.0	216	1 NOG1_BRARE	O9w741 brachydanio
42	31	62.0	218	1 SCOB_MYCTU	O06166 mycobacteri
43	31	62.0	222	1 NOGG_XENIA	P49011 xenopus lae
44	31	62.0	223	1 NOGG_BRARE	O9yhw3 brachydanio
45	31	62.0	223	1 NOGG_CHICK	O93525 gallus gall

ALIGNMENTS

RESULT 1
ID YHHT_ECOLI STANDARD: PRT: 349 AA.
AC P37622; P76700;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

GN YHHT OR B3474.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_Taxid=562;

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=94316500; PubMed=8041620;

RA Sofia H.-J., Burdand V., Daniels D.L., Plunkett G. III, Blattner F.R.;

RT "Analysis of the Escherichia coli genome. V. DNA sequence of the

region from 76.0 to 81.5 minutes."

RL Nucleic Acids Res. 22:2576-2586(1994).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE UPF0118 (PERK) FAMILY. STRONG, TO

H. INFLUENZAE H10338.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: 000039; CAB34657.1; ALT_INT.

CC EMBL: AE000423; AAC76499.1; ALT_INT.

DR Ecogene; EGI22220; YHHT.

DR InterPro; IPR002549; UPF0118.

DR Pfam; PF01594; UPF0118; 1.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 33 53 POTENTIAL.

FT TRANSMEM 63 83 POTENTIAL.

FT TRANSMEM 143 163 POTENTIAL.

FT TRANSMEM 209 229 POTENTIAL.

FT TRANSMEM 235 255 POTENTIAL.

FT TRANSMEM 258 278 POTENTIAL.

FT TRANSMEM 293 313 POTENTIAL.

SO SEQUENCE 349 AA; 38522 MW; 275ADBA04FD58257 CRC64;

Query Match 82.0%; Score 41; DB 1; Length 349;

Best Local Similarity 100.0%; Pred. No. 0.75;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSVC 9
|||||||

or send an email to license@lsb-slb.ch.

CC EMBL: K01544; AAA33005.1; -

DR PIR: A01330; NMRP1.

DR InterPro: IPR003612; AAI.

DR InterPro: IPR001768; Cereals_1tryp_1inh.

DR InterPro: IPR000617; Napin.

DR Pfam: PF00234; tryp_alpha_1.

DR Prodom: PD002498; Napin; 1.

DR SMART: SM00499; AAI; 1.

KW Seed storage protein; Multigene family.

FT NON_TER 1 1

FT CHAIN <1 30 SMALL CHAIN.

FT PROPEP 31 49

FT CHAIN 50 130 LARGE CHAIN.

SO SEQUENCE 133 AA; 15294 MW; 5FDC7E3E3E22ACB CRC64;

Query Match

Best Local Similarity 70.0%; Score 35; DB 1; Length 133;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPITSV 9

DB 114 CNIPQVSV 122

RESULT 5

2SS2_BRANA STANDARD; PRT; 178 AA.

AC P01090;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-AUG-1992 (Rel. 23, Last annotation update)

DE Napin 2 precursor (1.7S seed storage protein).

OS Brassica napus (Rape).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI_TaxID=3708;

RN [1]

RP SEQUENCE FROM N.A.

RA Josefsson L.-G.;

RL Submitted (JUL-1987) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=87033665; PubMed=3771543;

RA Ericson M.L.; Roedin J.; Lemman M.; Glimelius K.; Josefsson L.-G.;

RT "Structure of the rapeseed 1.7 S storage protein, napin, and its precursor."

RT J. Biol. Chem. 261:14576-14581(1986).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. TOWER;

RX MEDLINE=84113267; PubMed=6689334;

RA Crouch M.L.; Tenberge K.M.; Simon A.E.; Ferl R.;

RT "cDNA clones for Brassica napus seed storage proteins: evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide."

RT J. Mol. Appl. Genet. 2:273-283(1983).

RN [1]

RP FUNCTION: THE SMALL, BASIC, WATER-SOLUBLE NAPINS ARE ONE OF THE TWO MAJOR KINDS OF STORAGE PROTEINS SYNTHESIZED IN THE SEED DURING ITS MATURATION.

CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN LINKED BY DISULFIDE BONDS.

CC -1- TISSUE SPECIFICITY: COTYLEDONS AND THE AXIS.

CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.lsb-slb.ch/announce/> or send an email to license@lsb-slb.ch).

CC EMBL: K01545; AAA33006.1; -

DR EMBL: J02586; AAA32997.1; -

DR EMBL: J02798; AAA87348.1; -

DR PIR: A01329; NMRP2.

DR PIR: A29801; A29801.

DR PIR: A25997; A25997.

DR InterPro: IPR003612; AAI.

DR InterPro: IPR001768; Cereals_1tryp_1inh.

DR InterPro: IPR000617; Napin.

DR Pfam: PF00234; tryp_alpha_1.

DR PRINTS: PR00496; NAPIN.

DR Prodom: PD002498; Napin; 1.

DR SMART: SM00499; AAI; 1.

KW Seed storage protein; Signal; Multigene family.

FT SIGNAL 1 21

FT PROPEP 22 38 SMALL CHAIN.

FT CHAIN 39 74

FT PROPEP 75 94

FT CHAIN 95 175 LARGE CHAIN.

FT CONFLICT 37 37 D -> N (IN REF. 4).

FT CONFLICT 76 76 S -> N (IN REF. 4).

SO SEQUENCE 178 AA; 20104 MW; 734E561971B539FF CRC64;

Query Match

Best Local Similarity 70.0%; Score 35; DB 1; Length 178;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPITSV 9

DB 159 CNIPQVSV 167

RESULT 6

2SS3_BRANA STANDARD; PRT; 180 AA.

AC P17333;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Napin precursor (1.7S seed storage protein).

GN NAP1.

OS Brassica napus (Rape).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI_TaxID=3708;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. WESTAR; TISSUE=leaf;

RX MEDLINE=91346654; PubMed=2102844;

RA Baszczyński C.L.; Fallis L.;

RT "Isolation and nucleotide sequence of a genomic clone encoding a new Brassica napus napin gene."

RT Plant Mol. Biol. 14:633-635(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. TOPAS;

RX MEDLINE=95161697; PubMed=7858212;

RA Boutiller K.A.; Gines M.J.; Demoor J.M.; Huang B.;

RA Baszczyński C.L.; Iyer V.N.; Miki B.L.;

RT "Expression of the Bnmnap subfamily of napin genes coincides with the

RT Induction of Brassica microspore embryogenesis.";
 RL Plant Mol. Biol. 26:1711-1723(1994).
 CC -1- FUNCTION: THE SMALL, BASIC, WATER-SOLUBLE NAPINS ARE ONE OF THE
 CC TWO MAJOR KINDS OF STORAGE PROTEINS SYNTHESIZED IN THE SEED DURING
 CC ITS MATURATION.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- TISSUE SPECIFICITY: COTYLEDONS AND THE AXIS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17542; CAA35580.1; -;
 DR EMBL: U04945; AAA81909.1; -;
 DR PIR: S10018; S10018.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1-trypanin.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF00234; tryp_alpha_1; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 DR Seed storage protein; Signal; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 38
 FT CHAIN 39 74 SMALL CHAIN.
 FT PROPEP 75 94
 FT CHAIN 95 180 LARGE CHAIN.
 SQ SEQUENCE 180 AA; 20318 MW; 6F3883CBED55FB26 CRC64;
 Query Match 70.0%; Score 35; DB 1; Length 180;
 Best Local Similarity 55.6%; Pred. No. 5.9;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CSVPITSYVC 9
 Db 161 CNIPQVSVC 169
 RESULT 7
 TYRO_NEUCR STANDARD; PRT; 620 AA.
 AC P00440;
 DT 21-JUL-1986 (Rel. 01; Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Tyrosinase precursor (EC 1.14.18.1) (Monophenol monooxygenase).
 GN T.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariata; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=75, AND OR;
 RX MEDLINE=9000884; PubMed=2529259;
 RA Kupper U., Niedermann D.M., Travaglini G., Lerch K.;
 RT Isolation and characterization of the tyrosinase gene from
 RT Neurospora crassa.";
 RL J. Biol. Chem. 264:17250-17258(1989).
 RN [2]
 RP SEQUENCE OF 1-407.
 RC STRAIN=TL;
 RX MEDLINE=82190018; PubMed=6210696;
 RA Lerch K.;
 RT Primary structure of tyrosinase from Neurospora crassa. II. Complete

RT amino acid sequence and chemical structure of a tripeptide containing
 RT an unusual thioether.";
 RL J. Biol. Chem. 257:6414-6419(1982).
 RN [3]
 RP SEQUENCE OF 1-407.
 RC STRAIN=75, AND SING;
 RX MEDLINE=82190019; PubMed=6210697;
 RA Ruegg C., Ammer D., Lerch K.;
 RT "Comparison of amino acid sequence and thermostability of tyrosinase
 RT from three wild type strains of Neurospora crassa.";
 RL J. Biol. Chem. 257:6420-6426(1982).
 CC -1- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
 CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
 CC COMPOUNDS.
 CC -1- CATALYTIC ACTIVITY: L-tyrosine + L-DOPA + O(2) = L-DOPA +
 CC DOPAC + H(2)O.
 CC -1- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
 CC -1- MISCELLANEOUS: THE STRAIN 75 SEQUENCE IS SHOWN.
 CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M32843; AAA33619.1; -;
 DR EMBL: M3271; AAA33618.1; -;
 DR PIR: A00511; YRNC.
 DR PIR: A34460; A34460.
 DR InterPro: IPR002227; Tyrosinase.
 DR Pfam: PF00264; tyrosinase; 1.
 DR PRINTS: PR00092; TYROSINASE.
 DR PROSITE: PS00487; TYROSINASE_1; 1.
 DR PROSITE: PS00498; TYROSINASE_2; 1.
 DR Melanin biosynthesis; Oxidoreductase; Monooxygenase; Copper;
 KW acetylation.
 FT CHAIN 1 407
 FT INIT MET 0
 FT PROPEP 408 620
 FT MOD_RES 1 1
 FT METAL 66 66
 FT METAL 96 96
 FT METAL 105 105
 FT METAL 277 277
 FT METAL 281 281
 FT METAL 306 306
 FT THIOETH 94 96
 FT VARIANT 14 14
 FT VARIANT 29 29
 FT VARIANT 130 130
 FT VARIANT 345 346
 FT VARIANT 370 370
 FT VARIANT 423 423
 FT VARIANT 449 449
 FT CONFLICT 234 234
 SQ SEQUENCE 620 AA; 68546 MW; D12A30BB6A01D312 CRC64;
 Query Match 70.0%; Score 35; DB 1; Length 620;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SVPLTSYVC 9
 Db 589 TVPLTSLC 596
 RESULT 8
 ALL1_BRAU STANDARD; PRT; 129 AA.
 ID ALL1_BRAU

AC P80307; P80215;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Allergen Bra j 1-E, small and large chains (Bra j 1).
 OS Brassica juncea (leaf mustard) (Indian mustard).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Brassica.
 CC NCBI_TaxID=3707;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=93356721; PubMed=7688955;
 RA Monsalve R.I., Gonzalez de la Pena M.A., Menendez-Arias L.,
 RA Lopez-Otin C., Villalba M., Rodriguez R.;
 RT "Characterization of a new oriental-mustard (Brassica juncea)
 RL allergen, Bra j 1E: detection of an allergenic epitope.";
 CC Blochem. J. 293:625-632(1993).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY TWO DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC InterPro: IPR003612; AAI.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF01631; Seedstore_2S; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 KM Allergen; Seed storage protein.
 FT CHAIN 1 SMALL CHAIN.
 FT NON_CONS 37 38
 FT CHAIN 38 129 LARGE CHAIN.
 FT VARIANT 6 6 F -> I.
 FT VARIANT 20 20 R -> K.
 SQ SEQUENCE 129 AA; 14644 MW; D6F28E03F62B08F8 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 129;
 Best Local Similarity 44.4%; Pred. No. 6.7;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITVC 9
 DB 111 CNIPVSVIC 119

RESULT 9
 2SSB_BRANA STANDARD; PRT; 178 AA.
 ID 2SSB_BRANA P27740;
 AC 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Napin B precursor (1.75 seed storage protein).
 GN NABP.
 OS Brassica napus (Rape).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Brassica.
 CC NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SVALOF KARAT 20516-K;
 RX MEDLINE=91231016; PubMed=2029903;
 RA Ericson M.L., Muren E., Gustavsson H.O., Josefsson L.G., Rask L.;
 RT "Analysis of the promoter region of napin genes from Brassica napus
 RT demonstrates binding of nuclear protein in vitro to a conserved
 RT sequence motif.";
 CC Eur. J. Biochem. 197:741-746(1991).
 CC -1- FUNCTION: THE SMALL, BASIC, WATER-SOLUBLE NAPINS ARE ONE OF THE
 CC TWO MAJOR KINDS OF STORAGE PROTEINS SYNTHESIZED IN THE SEED DURING
 CC ITS MATURATION.

CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- TISSUE SPECIFICITY: COTYLEDONS AND THE AXIS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X58142; CAA41150.1; -
 DR PIR: S15382; S15382.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal-tryr-amy1_inh.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF00234; tryr-alpha_amy1; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 KM Seed storage protein; Signal; Multigene family; Embryo.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT PROPEP 22 38 BY SIMILARITY.
 FT CHAIN 39 74 SMALL CHAIN (BY SIMILARITY).
 FT PROPEP 75 94 BY SIMILARITY.
 FT CHAIN 95 178 LARGE CHAIN (BY SIMILARITY).
 SQ SEQUENCE 178 AA; 20114 MW; 96CE0ADB7CD966E9 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 178;
 Best Local Similarity 55.6%; Pred. No. 9.2;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSVPITVC 9
 DB 159 CKIPVSVIC 167

RESULT 10
 EPA2_RALSO STANDARD; PRT; 377 AA.
 ID EPA2_RALSO Q45407;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE EPS I polysaccharide export outer membrane protein epsa precursor.
 GN EPSA.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 CC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 CC Ralstonia.
 CC NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM.
 RX MEDLINE=96059643; PubMed=7476194;
 RA Huang J., Schell M.;
 RT "Molecular characterization of the eps gene cluster of Pseudomonas
 RT solanacearum and its transcriptional regulation at a single
 RT promoter.";
 CC Mol. Microbiol. 16:977-989(1995).
 CC -1- FUNCTION: PROBABLY INVOLVED IN POLYMERIZATION AND/OR EXPORT OF.
 CC EXOPOLYSACCHARIDE EPS I WHICH FUNCTIONS AS A VIRULENCE FACTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
 CC (potential).
 CC -1- SIMILARITY: BELONGS TO THE BEXD/CTRA/VEXA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U17898; AAA91624.1; -

DR InterPro: IPR003715; Poly_export.

DR InterPro: IPR000437; Prok_lipoprot.

DR Pfam: PF02563; Poly_export; 1.

DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; FALSE_NEG.

KW Polysaccharide transport; Transport; Outer membrane; Transmembrane;

KW Lipoprotein; Porin; Signal.

FT SIGNAL 1 23

FT CHAIN 24 377

FT POTENTIAL.

FT ERS I POLYSACCHARIDE EXPORT OUTER

FT MEMBRANE PROTEIN EPSA.

FT N-ACYL DIGLYCERIDE (BY SIMILARITY).

SO SEQUENCE 377 AA; 40446 MW; BE0084252E819309 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 377;

Best Local Similarity 55.6%; Pred. No. 20;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPITVVC 9

DB 16 CAVPLMAC 24

RESULT 11

PIZR_HUMAN STANDARD; PRT; 386 AA.

ID P3119;

AC 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Prostacyclin receptor (Prostanoid IP receptor) (PGI receptor).

GN PGIR OR PRIPR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=94216334; PubMed=7512962;

RA Bole Y., Rushmore T.H., Datmon-Goodwin A., Grygorczyk R.,

RA Sliet D.M., Metters K.M., Abramovitz M.;

RT "Cloning and expression of a cDNA for the human prostanoid IP

RT receptor.";

RL J. Biol. Chem. 269:12173-12178(1994).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94237286; PubMed=7514139;

RA Katsuyama M., Sugimoto Y., Namba T., Irie A., Negishi M.,

RA Natumiya S., Ichikawa A.;

RT "Cloning and expression of a cDNA for the human prostacyclin

RT receptor.";

RL FEBS Lett. 344:74-78(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=95008086; PubMed=7923647;

RA Nakagawa O., Tanaka I., Usui T., Harada M., Sasaki Y., Itoh H.,

RA Yoshimasa T., Namba T., Natumiya S., Nakao K.;

RT "Molecular cloning of human prostacyclin receptor cDNA and its gene

RT expression in the cardiovascular system.";

RL Circulation 90:1643-1647(1994).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=95394450; PubMed=7665161;

RA Ogawa Y., Tanaka I., Inoue M., Yoshitake Y., Ise N., Nakagawa O.,

RA Usui T., Itoh H., Yoshimasa T., Natumiya S.;

RT "Structural organization and chromosomal assignment of the human

RT prostacyclin receptor gene.";

RL Genomics 37:142-148(1995).

CC -1- FUNCTION: RECEPTOR FOR PROSTACYCLIN (PROSTAGLANDIN I2 OR PGI2).

CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G-S PROTEINS WHICH

CC ACTIVATE ADENYLYLATE CYCLASE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: L29016; AAA36448.1; -

DR EMBL: D25418; BAA05008.1; -

DR EMBL: D29634; BAA06110.1; -

DR EMBL: D38127; BAA07325.1; -

DR EMBL: D38128; BAA07325.1; JOINED.

DR GCRDB: GCR_0941; -

DR GCRDB: GCR_1159; -

DR GCRDB: GCR_1830; -

DR MIM: 600022; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS: PR00856; PRSTNOIDIPR.

DR PROSITE: PS00237; G-PROTEIN_REC_P1.1; 1.

DR PROSITE: PS50262; G-PROTEIN_REC_P1.2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 16

FT TRANSMEM 17 38

FT DOMAIN 39 51

FT TRANSMEM 52 76

FT DOMAIN 77 94

FT TRANSMEM 95 115

FT DOMAIN 116 134

FT TRANSMEM 135 158

FT DOMAIN 159 181

FT TRANSMEM 182 208

FT DOMAIN 209 235

FT TRANSMEM 236 260

FT DOMAIN 261 274

FT TRANSMEM 275 295

FT DOMAIN 296 386

FT CARBOHYD 7

SO SEQUENCE 386 AA; 40955 MW; 2B6B0CDBACE1D1608 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 386;

Best Local Similarity 66.7%; Pred. No. 20;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPITVVC 9

DB 251 CSPLPTRC 259

RESULT 12

CIKS_HUMAN STANDARD; PRT; 574 AA.

ID 043734; Q9UG72; Q9NS14;

AC 043734; Q9UG72; Q9NS14;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Adapter protein CIKS (Connection to IKK and SAPK/JNK).

GN C6ORF4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Liver;

RX MEDLINE-20442403; PubMed-10962033;
 RA Leonard A., Charlot A., Claudio E., Cunnigham K., Siebenlist U.;
 RT "Clk, a connection to Ikappa B kinase and stress-activated protein
 kinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10494-10499(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE-20363742; PubMed-10903453;
 RA Morelli C., Magnanini C., Mungall A.J., Negri M.,
 RA Baraniti-Brodano G.;
 RT "Cloning and characterization of two overlapping genes in a subregion
 at 6q21 involved in replicative senescence and schizophrenia.";
 RL Gene 252:217-225(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Palmer S.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-uterus;
 RA Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-lymph;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: COULD BE INVOLVED IN THE ACTIVATION OF BOTH NF-KAPPA-B
 CC VIA A NF-KAPPA-B INHIBITOR KINASE (IKK)-DEPENDENT MECHANISM AND
 CC STRESS-ACTIVATED PROTEIN KINASE (SAPK)/JNK.
 CC -1- SUBUNIT: INTERACTS WITH IKKKG/NF-KAPPA B ESSENTIAL MODULATOR, WITH
 CC CHUK/IKK-ALPHA AND WITH IKKKG/IKK-BETA.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/60RF4 (SHOWN HERE) AND
 CC 2/60RF5/60RF6; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF272151; AAG15407.1; -
 DR EMBL: AF136405; AAF67445.1; -
 DR EMBL: AF136406; AAF67446.1; -
 DR EMBL: AF136407; AAF67447.1; -
 DR EMBL: AL008730; CA15506.1; -
 DR EMBL: AL008730; CA15507.1; ALT_INIT.
 DR EMBL: AL050289; CAB43390.1; -
 DR EMBL: BC002823; AAH02823.1; -
 KW Alternative splicing.
 KM VASPLIC 1
 FT CONFLICT 332 332 MISSING (IN ISOFORM 2).
 FT CONFLICT 334 334 E -> D (IN REF. 1) AAF67445).
 FT CONFLICT 347 347 P -> S (IN REF. 2) AAF67447).
 SQ SEQUENCE 574 AA; 64657 MW; E9857DEA5E349094 CRC64;
 QY 3 VPLTSVC 9
 DB 225 LPLTSVC 231
 Query Match 68.0%; Score 34; DB 1; Length 574;
 Best Local Similarity 85.7%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ARGRS).
 GN RARS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96069607; PubMed-7590355;
 RA Girjes A.A., Hobson K., Chen P., Lavin M.F.;
 RT "Cloning and characterization of cDNA encoding a human arginyl-tRNA
 RT synthetase.";
 RL Gene 164:347-350(1995).
 CC -1- CATALYTIC ACTIVITY: Arg + L-arginine + tRNA(Arg) -> AMP +
 CC di-phosphate + L-arginyl-tRNA(Arg).
 CC -1- SUBUNIT: MONOMER; ALSO PART OF A MULTISUBUNIT COMPLEX THAT GROUPS
 CC tRNA LIGASES FOR ARG, ASP, GLU, GLN, ILE, LEU, LYS, MET AND PRO.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DOMAIN: THE N-TERMINAL (AA 1-72) HAS TWO REGIONS PREDICTED TO BE
 CC ALPHA-HELICAL THAT MIGHT BE INVOLVED IN THE MULTISUBUNIT
 CC COMPLEX ASSEMBLY.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S80343; AAB35627.1; -
 DR MIM: 107820; -
 DR InterPro: IPR001278; tRNA-synt_1d.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00750; tRNA-synt_1d; 1.
 DR PRINTS: PR01038; TRNASYNTARG.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Alternative initiation.
 KM CHAIN 1 659
 FT FT
 FT CHAIN 73 659
 FT FT
 FT INIT_MET 73 73
 FT DOMAIN 1 72
 FT FT
 FT SITE 201 212
 SQ SEQUENCE 659 AA; 74977 MW; 91FBL7F9E04AC3D9 CRC64;
 QY 1 CVPVTSV 8
 DB 369 CSIPPLTV 376
 Query Match 68.0%; Score 34; DB 1; Length 659;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 14
 ID SYR_CRILLO
 AC P37880; STANDARD; PRT; 661 AA.
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ARGRS).
 GN RARS OR RRS1.
 OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Criceulius.
 OX NCBI_TaxID=10030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=94040767; PubMed=8224869;
 RA Lazard M., Mirande M.;
 RT "Cloning and analysis of a cDNA encoding mammalian arginyl-tRNA
 RT synthetase, a component of the multisynthetase complex with a
 RT hydrophobic N-terminal extension.";
 RL Gene 132:237-245(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) - AMP +
 CC diposphate + L-arginyl-tRNA(Arg).
 CC -1- SUBUNIT: MONOMER; ALSO PART OF A MULTISUBUNIT COMPLEX THAT GROUPS
 CC TRNA LIGASES FOR ARG, ASP, GLU, GLN, ILE, LEU, LYS, MET AND PRO.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: THE USE OF ALTERNATIVE INITIATION CODONS IN
 CC THE SAME READING FRAME MAY GIVE RISE TO A COMPLEXED AND A FREE
 CC MONOMERIC SPECIES.
 CC -1- DOMAIN: THE N-TERMINAL (AA 1-73) HAS TWO REGIONS PREDICTED TO BE
 CC ALPHA-HELICAL THAT MIGHT BE INVOLVED IN THE MULTISYNTHETASE
 CC COMPLEX ASSEMBLY.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X63415; CAA5012.1; -
 DR PIR: JN0870; JN0870.
 DR InterPro: IPR001278; trna-synt_1d.
 DR InterPro: IPR001412; trna-synt_1.
 DR Pfam: PF00750; trna-synt_1d; 1.
 DR PRINTS: PR01038; TRNASYNTHARG.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Alternative initiation.
 FT CHAIN 1 661 ARGINYL-TRNA SYNTHETASE, COMPLEXED
 FT ISOFORM.
 FT CHAIN 74 661 ARGINYL-TRNA SYNTHETASE, MONOMERIC
 FT INIT_MET 74 74 ISOFORM (POTENTIAL)
 FT DOMAIN 1 73 FOR MONOMERIC ISOFORM (POTENTIAL).
 FT COULD BE INVOLVED IN THE ASSEMBLY OF THE
 FT MULTISYNTHETASE COMPLEX.
 FT SITE 202 213 "HIGH" REGION.
 FT SEQUENCE 661 AA; 75602 MW; 12EB1C85655EB8F4 CRC64;
 SQ
 Query Match 68.0%; Score 34; DB 1; Length 661;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OS Rickettsia conorii.
 CC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissensbach U., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 CC -1- FUNCTION: Furnishes a means for formation of correctly charged
 CC Gln-tRNA(Gln) through the transamidation of misacylated Glu-
 CC tRNA(Gln) in organisms which lack glutamyl-tRNA synthetase. The
 CC reaction takes place in the presence of glutamine and ATP through
 CC an activated gamma-phospho-Glu-tRNA(Gln) (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Gln) + L-glutamine - ADP
 CC + phosphate + L-glutamyl-tRNA(Gln) + L-glutamate.
 CC -1- SUBUNIT: Heterotrimer of A, B and C subunits (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE GATC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE00587; AL02733.1; ALT_INIT.
 DR Protein biosynthesis; Ligase; Complete proteome.
 KW SEQUENCE 100 AA; 11321 MW; 41BF2C04E7346100 CRC64;
 SQ
 Query Match 66.0%; Score 33; DB 1; Length 100;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PLTSVC 9
 DB 50 PLTSVC 55
 Search completed: June 17, 2002, 16:24:29
 Job time: 1558 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:39 ; Search time 172.85 Seconds
(without alignments)
9.008 Million cell updates/sec

Title: US-09-761-636A-10

Perfect score: 50

Sequence: 1 CSVPRTSVC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oranelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	86.0	1487	11	062028 mus musculu
2	41	82.0	409	11	063791 rattus sp.
3	41	82.0	750	10	09ARC9 lycopersico
4	41	82.0	750	10	09ARC6 lycopersico
5	41	82.0	780	10	09ARF2 lycopersico
6	41	82.0	784	10	09ARF3 arabidopsis
7	41	82.0	784	10	09LGM4 arabidopsis
8	39	78.0	432	17	058874 methanococ
9	38	76.0	491	4	09H6R1 homo sapien
10	38	76.0	786	4	09H747 homo sapien
11	37	74.0	718	10	094JF5 oryza sativ
12	37	74.0	1929	5	09V6A0 drosophila
13	36	72.0	124	11	09D179 mus musculu
14	36	72.0	168	4	096PA7 homo sapien
15	36	72.0	201	4	09H1K3 homo sapien
16	36	72.0	247	4	09BXN2 homo sapien

17	36	72.0	591	10	039485 chlamydomon
18	36	72.0	1217	4	09ULB5 homo sapien
19	36	72.0	2192	5	001768 caenorhabdi
20	35	70.0	86	10	09S9F0 brassica na
21	35	70.0	88	10	09S9E9 brassica na
22	35	70.0	88	10	09S9E8 brassica na
23	35	70.0	88	10	09S9E7 brassica na
24	35	70.0	91	16	097HY9 clostridium
25	35	70.0	125	10	P80208 brassica na
26	35	70.0	145	10	041281 sinapis alb
27	35	70.0	145	10	041278 sinapis alb
28	35	70.0	145	10	041279 sinapis alb
29	35	70.0	145	10	041280 sinapis alb
30	35	70.0	152	2	09ZB10 streptomyce
31	35	70.0	155	10	041167 raphanus sa
32	35	70.0	168	5	09NL63 sarcophaga
33	35	70.0	178	10	042413 brassica ju
34	35	70.0	178	10	042490 brassica ol
35	35	70.0	178	10	042491 brassica ni
36	35	70.0	178	10	039344 brassica na
37	35	70.0	178	10	096339 brassica ca
38	35	70.0	180	10	042469 brassica na
39	35	70.0	185	6	09MZ36 macaca mula
40	35	70.0	189	4	043198 homo sapien
41	35	70.0	216	6	09MZ37 macaca mula
42	35	70.0	327	4	076091 homo sapien
43	35	70.0	378	10	09LHM1 oryza sativ
44	35	70.0	660	11	09D019 mus musculu
45	34	68.0	91	10	09S9E5 brassica na

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	1487 AA.
062028	AC	062028;		
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	PHOSPHOLIPASE A2 RECEPTOR PRECURSOR.			
GN	PLA2G1BR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]	SEQUENCE FROM N.A.		
RP	MEDLINE=95010128; PubMed=7925459;			
RX	Higashino K., Ishizaki J., Kishino J., Ohara O., Arita H.;			
RA	"Structural comparison of phospholipase-A2-binding regions in			
RT	phospholipase-A2 receptors from various mammals."			
RL	Eur. J. Biochem. 225:375-382(1994).			
DR	EMBL: D30779; BAA06443.1; -.			
DR	HSSP: P02751; 2FN2.			
DR	MGI: 102468; Pla2g1br.			
DR	InterPro: IPR000562; FN_Type_II.			
DR	InterPro: IPR001304; lectin_c.			
DR	InterPro: IPR000772; Ricin_B_lectin.			
DR	Pfam: PF00407; In2; 1.			
DR	Pfam: PF00059; lectin_c; 8.			
DR	Pfam: PF00652; Ricin_B_lectin; 1.			
DR	PRINTS: PR00013; FNTYPEII.			
DR	PRODOM: PD000995; FN_Type_II; 1.			
DR	SMART: SM00034; CLECT; 8.			
DR	SMART: SM00059; FN2; 1.			
DR	SMART: SM00458; RICIN; 1.			
DR	PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_2.			
DR	PROSITE: PS50041; C_TYPE_LECTIN_2; 8.			
DR	PROSITE: PS00023; FIBRONECTIN_2; UNKNOWN_1.			
DR	PROSITE: PS50231; RICIN_B_LECTIN; 1.			
KW	Signal; Receptor.			

FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1487 PHOSPHOLIPASE A2 RECEPTOR.
 SQ SEQUENCE 1487 AA; 170511 MW; AD8D905859B0EDE8 CRC64;

Query Match
 Best Local Similarity 86.0%; Score 43; DB 11; Length 1487;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSQLPLTVC 9
 |||||111
 Db 929 CSQLPLTVC 937

RESULT 2

063791 PRELIMINARY; PRT: 409 AA.

AC 063791; 01-NOV-1996 (TREMBLERL. 01, Created)
 DT 01-NOV-1996 (TREMBLERL. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBLERL. 19, Last annotation update)
 OS PHOSPHOLIPASE A2 RECEPTOR (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMOOTH MUSCLE;
 RX MEDLINE=95010128; PubMed=7925459;
 RA Higashino K., Ishizaki J., Kishino J., Ohara O., Arita H.;
 RT "Structural comparison of phospholipase-A2-binding regions in
 RL Eur. J. Biochem. 225:375-382(1994).
 DR EMBL: D30781; BAA06445.1; -.
 DR HSP: P23807; I1XX.
 DR InterPro: IPR001304; Lectin_c.
 DR Pfam: PF00059; Lectin_c; 3.
 DR SMART: SM00034; CLECT; 3.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 3.
 KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 409 409
 SQ SEQUENCE 409 AA; 47740 MW; 56D957D2DCA0A0AD8 CRC64;

Query Match
 Best Local Similarity 82.0%; Score 41; DB 11; Length 409;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSQLPLTVC 9
 |||||111
 Db 396 CSQLPLTVC 404

RESULT 3

09ARC9 PRELIMINARY; PRT: 750 AA.

AC 09ARC9; 01-JUN-2001 (TREMBLERL. 17, Created)
 DT 01-JUN-2001 (TREMBLERL. 17, Last sequence update)
 DE 01-OCT-2001 (TREMBLERL. 18, Last annotation update)
 OS HYPOTHETICAL 80.0 KDA PROTEIN.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21178822; PubMed=11283350;
 RA Rosenberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,
 Schumacher K., Schmitt G., Schmidt R.;

RT "Comparative sequence analysis reveals extensive microcolinearity in
 RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella
 RT genomes.";
 RL Plant Cell 13:979-988(2001).
 DR EMBL: AJ303344; CAC36398.1; -.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR01807; Volt_Cl_channel.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; voltage_CLC; 1.
 DR PRINTS: PR00762; CLCHANNEL.
 DR SMART: SM00116; CBS; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 750 AA; 79994 MW; 400448BE3D37042C CRC64;

QY 1 CSQLPLTVC 8
 |||||111
 Db 474 CSQLPLTVC 481

Query Match
 Best Local Similarity 82.0%; Score 41; DB 10; Length 750;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSQLPLTVC 8
 |||||111
 Db 474 CSQLPLTVC 481

RESULT 4

09ARC6 PRELIMINARY; PRT: 750 AA.

AC 09ARC6; 01-JUN-2001 (TREMBLERL. 17, Created)
 DT 01-JUN-2001 (TREMBLERL. 17, Last sequence update)
 DE 01-OCT-2001 (TREMBLERL. 18, Last annotation update)
 OS HYPOTHETICAL 80.0 KDA PROTEIN.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21178822; PubMed=11283350;
 RA Rosenberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,
 Schumacher K., Schmitt G., Schmidt R.;
 RT "Comparative sequence analysis reveals extensive microcolinearity in
 RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella
 RT genomes.";
 RL Plant Cell 13:979-988(2001).
 DR EMBL: AJ303345; CAC36403.1; -.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR01807; Volt_Cl_channel.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; voltage_CLC; 1.
 DR PRINTS: PR00762; CLCHANNEL.
 DR SMART: SM00116; CBS; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 750 AA; 79985 MW; 928E6D1EE25FCD4 CRC64;

Query Match
 Best Local Similarity 82.0%; Score 41; DB 10; Length 750;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSQLPLTVC 8
 |||||111
 Db 474 CSQLPLTVC 481

RESULT 5

09ARF2 PRELIMINARY; PRT: 780 AA.

AC 09ARF2; 01-JUN-2001 (TREMBLERL. 17, Created)
 DT 01-JUN-2001 (TREMBLERL. 17, Last sequence update)
 DE 01-OCT-2001 (TREMBLERL. 18, Last annotation update)

DE HYPOTHETICAL 83.8 KDA PROTEIN.
 OS Capsella rubella.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Capsella.
 OX NCBI_TaxID=81985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21178822; PubMed-11283350;
 RA Rossberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,
 RA Schumacher K., Schmitt G., Schmidt R.;
 RT "Comparative sequence analysis reveals extensive microcolinearity in
 RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella
 RT genomes.";
 RL Plant Cell 13:979-988(2001).
 DR EMBL: AJ303349; CAC36391.1; -.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR001807; Volt_CL_channel.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; voltage_CIC; 1.
 DR PRINTS: PR00762; CICHANNEL.
 DR SMART: SM00116; CBS; 2.
 DR Hypothetical protein.
 KW SEQUENCE 780 AA; 83811 MW; 5718FA7E2AE81FB CRC64;

Query Match 82.0%; Score 41; DB 10; Length 780;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSVPILTSV 8
 |||||
 DB 513 CSVPILTSV 520

RESULT 6
 O9ARM3 PRELIMINARY; PRT; 781 AA.
 ID O9ARM3;
 AC O9ARM3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE HYPOTHETICAL 83.5 KDA PROTEIN (CIC-F CHLORIDE CHANNEL PROTEIN).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21178822; PubMed-11283350;
 RA Rossberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,
 RA Schumacher K., Schmitt G., Schmidt R.;
 RT "Comparative sequence analysis reveals extensive microcolinearity in
 RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella
 RT genomes.";
 RL Plant Cell 13:979-988(2001).
 DR EMBL: AJ303349; CAC36391.1; -.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR001807; Volt_CL_channel.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; voltage_CIC; 1.
 DR PRINTS: PR00762; CICHANNEL.
 DR SMART: SM00116; CBS; 2.

KW Hypothetical protein.
 OS SEQUENCE 781 AA; 83548 MW; BE9DEB3603D9E0D8 CRC64;

Query Match 82.0%; Score 41; DB 10; Length 781;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSVPILTSV 8
 |||||
 DB 514 CSVPILTSV 521

RESULT 7
 O9LG04 PRELIMINARY; PRT; 784 AA.
 ID O9LG04;
 AC O9LG04;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE F20N2.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
 RA Shin P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu K., Liu S., Mukharasy N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
 RT 1.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Shin P.P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,
 RA Walker M.M., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
 RA Gonzalez A.A., Hansen N.N.F., Huizar L.L., Kremetskaia I.I.,
 RA Lenz C.C., Li J.J., Liu S.S., Luos S.S., Rowley D.D., Schwartz J.J.,
 RA Toriumi M.M., Vysotskaia V.V., Yu G.G., Davis R.R.W.,
 RA Federspiel N.N.A., Theologis A.A., Ecker J.J.R.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharasy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002328; AAF79509.1; -.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR001807; Volt_CL_channel.
 DR Pfam: PF00571; CBS; 2.

DR Pfam: PF00654; voltage-CLC; 1.
 DR PRINTS: PR00762; CLCHANNEL.
 DR SMART: SM00116; CBS; 2.
 SQ SEQUENCE 784 AA; 83614 MW; 0495177BDAL9330D CRC64;

Query Match
 Best Local Similarity 82.0%; Score 41; DB 10; Length 784;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CSVPITSV 8
 Db 514 CSVPITSV 521

RESULT 8
 ID 058874 PRELIMINARY; PRT; 432 AA.
 AC 058874;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHEICAL PROTEIN MJ1479.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-JUL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-96337999; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.V., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reick C.I.,
 RA Overbeek R., Kirkness E.F., Weisslock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uitterlank T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: TO E. COLI HYPOTHEICAL 45.5 KDA PROTEIN IN LRHA
 3-REGION. H. INFLUENZAE HI0286 AND M. TUBERCULOSIS MT0279.04C.
 DR EMBL: U67589; AAB9491.1; -;
 DR TIGR: MJ1479; -;
 DR InterPro: IPR001511; Aminoctran_1.
 DR InterPro: IPR000529; Ribosomal_S6.
 DR Pfam: PF00155; aminotran_1.2; 1.
 DR PROSITE: PS01048; RIBOSOMAL_S6; UNKNOWN_1.
 KW Hypothetical protein; complete proteome.
 SQ SEQUENCE 432 AA; 49429 MW; 14D1602E32DB31FE CRC64;

Query Match
 Best Local Similarity 78.0%; Score 39; DB 17; Length 432;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITSV 9
 Db 389 CSVPITSV 397
 RESULT 9
 ID 09H6R1 PRELIMINARY; PRT; 491 AA.
 AC 09H6R1;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CNA: FLJ21969 FIS, CLONE HEP05692.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK025622; BAB15192.1; -;
 SQ SEQUENCE 491 AA; 54720 MW; 4F73721A9D092C37 CRC64;

Query Match
 Best Local Similarity 76.0%; Score 38; DB 4; Length 491;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITSV 9
 Db 475 CSVPITDPC 483

RESULT 10
 ID 09H747 PRELIMINARY; PRT; 786 AA.
 AC 09H747;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CDNA: FLJ21347 FIS, CLONE COL02724.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK025000; BAB15051.1; -;
 DR InterPro: IPR000834; Zn_cardioph.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 SQ SEQUENCE 786 AA; 87893 MW; 958D17DD68B321C2 CRC64;

Query Match
 Best Local Similarity 76.0%; Score 38; DB 4; Length 786;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITSV 9
 Db 770 CSVPITDPC 778

RESULT 11
 ID 094JF5 PRELIMINARY; PRT; 718 AA.
 AC 094JF5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE P0684B02.13 PROTEIN.
 GN P0684B02.13.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0684B02."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003023; BAB44026.1;
 SQ SEQUENCE 718 AA; 76362 MW; DAB54F202012785B CRC64;

Query Match 74.0%; Score 37; DB 10; Length 718;
 Best Local Similarity 87.5%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSVPRTSV 8
 Db 463 CWPVLTSTV 470

RESULT 12
 ID 09V6A0 PRELIMINARY; PRT; 1929 AA.
 AC 09V6A0:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE CG13162 PROTEIN.
 GN CG13162.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
 Abril J.F., Abdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Duran-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaitani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstock G.M., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.C.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003822; AAF58528.1;
 DR Flybase: FBgn0033718; CG13162.
 SQ SEQUENCE 1929 AA; 219961 MW; A409187CC6AECAC66 CRC64;

Query Match 74.0%; Score 37; DB 5; Length 1929;
 Best Local Similarity 55.6%; Pred. No. 83;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSVPRTSV 9
 Db 1327 CSIPVETLC 1335

RESULT 13
 ID 09D179 PRELIMINARY; PRT; 124 AA.
 AC 09D179:
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 110020C17RIK PROTEIN.
 GN 110020C17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 Fleischman W., Gaasterland T., Gissi C., King B., Kochiya H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quekenuush J.,
 Schirrali L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gastincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK003851; BAB23036.1;
 DR MGD: MGI:1915834; 110020C17RIK.
 SQ SEQUENCE 124 AA; 13504 MW; 804010C4E2C4F75 CRC64;

Query Match 72.0%; Score 36; DB 11; Length 124;
 Best Local Similarity 55.6%; Pred. No. 11;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSVPRTSV 9
 Db 56 CWPVTEMC 64

RESULT 14
 ID 096PA7 PRELIMINARY; PRT; 168 AA.
 AC 096PA7:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE BETA-GLUCAN RECEPTOR ISOFORM E.
 GN BGR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Willment J.A., Gordon S., Brown G.D.;
 RT "Characterization of the human beta-glucan receptor and its
 RT alternatively spliced isoforms.";
 RL J. Biol. Chem. 0:0-0(2001).
 DR EMBL: AF400599; AAL11715.1; -
 KW Receptor.
 SQ SEQUENCE 168 AA; 19217 MW; AFD3A8BF89BFFC6 CRC64;

 Query Match 72.0%; Score 36; DB 4; Length 168;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 1 CSVP1RSVC 9
 DB 154 CSVP1RSVC 162

 RESULT 15
 ID Q9H1K3 PRELIMINARY; PRT; 201 AA.
 AC Q9H1K3;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE TRANSMEMBRANE PROTEIN DECTIN-1 (DENDRITIC CELL-ASSOCIATED C-
 DE TYPE LECTIN-1 BETA) (DECTIN-1 RECEPTOR) (LECTIN-LIKE RECEPTOR 1B)
 DE (BETA-GLUCAN RECEPTOR ISOFORM B).
 GN DECTIN-1 OR DECTIN1 OR BGR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Grunebach F., Brugger W., Kanz L., Brossart P.;
 RT "Identification of the human dendritic-cell-associated molecule,
 RT hdectin-1, by cDNA subtraction and Rapid Amplification of cDNA Ends
 RT (RACE).";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yokota K., Takashima A., Bergstresser P.R., Arizumi K.;
 RT "Identification of a Human Homolog of the Dendritic Cell-associated C-
 RT type Lectin-1, Dectin-1.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Sobanov Y., Bernreiter A., Derdak S., Mechtyerlakova D., Duechler M.,
 RA Kalthoff F., Hofer E.;
 RT "A novel cluster of lectin-like receptor genes expressed in monocytic,
 RT dendritic and endothelial cells maps close to the NK receptor genes in
 RT the human NK gene complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21383615; Pubmed-11491532;
 RA Hernandez-Falcon P., Arce I., Roda Navarrio P., Fernandez-Ruiz E.;
 RT "Cloning of human DECTIN-1, a novel C-type lectin-like receptor gene
 RT expressed on dendritic cells.";
 RL Immunogenetics 53:288-295(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Hernandez-Falcon P., Arce I., Fernandez-Ruiz E.;

RL submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PERIPHERAL BLOOD LEUKOCYTE;
 RA Willment J.A., Gordon S., Brown G.D.;
 RT "Characterization of the human beta-glucan receptor and its
 RT alternatively spliced isoforms.";
 RL J. Biol. Chem. 0:0-0(2001).
 DR EMBL: AY009090; AAG33923.2; -
 DR EMBL: AF313469; AAK37474.1; -
 DR EMBL: AJ312372; CAC43846.1; -
 DR EMBL: AY026770; AAK20115.1; -
 DR EMBL: AF400596; AAL11712.1; -
 DR InterPro: IPR002353; AntifreezeII.
 DR InterPro: IPR001304; lectin_c.
 DR Pfam: PF00059; lectin_c.1.
 DR PRINTS: PR00356; ANTIFREEZEII.
 DR SMART: SM00034; CLECT.1.
 DR PROSITE: PS50041; C_Type_Lectin_2; 1.
 KW Transmembrane; Lectin; Receptor.
 SQ SEQUENCE 201 AA; 22563 MW; C6ADEE762B2CE968 CRC64;

 Query Match 72.0%; Score 36; DB 4; Length 201;
 Best Local Similarity 66.7%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 1 CSVP1RSVC 9
 DB 187 CSVP1RSVC 195

Search completed: June 17, 2002, 16:23:40
 Job time: 1614 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:13 ; Search time 227.35 Seconds
(without alignments)
4.397 Million cell updates/sec

Title: US-09-761-636a-10
Perfect score: 50
Sequence: 1 CSVPYRVSVC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	100.0	9	22	AAU04529	VEGF based monocyc
2	43	86.0	1487	21	AA23826	Murine phospholipa
3	39	78.0	73	22	AAU63687	Protonibacterium
4	39	78.0	88	22	AAU53277	Protonibacterium
5	38	76.0	101	22	AAU43894	Protonibacterium
6	38	76.0	225	22	ABG11475	Novel human diagno
7	38	76.0	742	22	AA27463	Human protease and
8	38	76.0	802	22	AAE01782	Human gene 13 enco
9	37	74.0	86	22	AAU58469	Protonibacterium
10	37	74.0	1929	22	ABB67846	Drosophila melanog
11	36	72.0	90	22	AA674036	Human colon cancer

12 36 72.0 133 22 ABB10143
13 36 72.0 133 22 AAU18408
14 36 72.0 133 22 AAM43498
15 36 72.0 133 22 AAU19674
16 36 72.0 133 22 AAU21569
17 36 72.0 201 19 AAM52837
18 36 72.0 201 20 AAU27449
19 36 72.0 201 20 AAU41764
20 36 72.0 201 20 AAM73888
21 36 72.0 201 21 AAB44320
22 36 72.0 201 22 AAU29077
23 36 72.0 247 20 AAU27448
24 36 72.0 247 20 AAM73889
25 36 72.0 489 15 AAR63377
26 36 72.0 900 21 AAB42321
27 36 72.0 1217 22 AAG09876
28 35 70.0 54 22 AAM06633
29 35 70.0 77 22 ABB28805
30 35 70.0 77 22 ABB33992
31 35 70.0 77 22 ABB19430
32 35 70.0 77 22 AAM64755
33 35 70.0 77 22 AAM67144
34 35 70.0 77 22 AAM15009
35 35 70.0 77 22 AAM27446
36 35 70.0 77 22 AAM02737
37 35 70.0 137 21 AAB10661
38 35 70.0 137 21 AAB10663
39 35 70.0 157 22 AAU57063
40 35 70.0 175 14 AAR33390
41 35 70.0 189 20 AAM88255
42 35 70.0 189 20 AAE11759
43 35 70.0 224 21 AAB57054
44 35 70.0 464 21 AAY68739
45 34.5 69.0 59 22 AAU66174

ALIGNMENTS

RESULT 1
ID AAU04529 standard; Peptide: 9 AA.
XX
AC AAU04529;
XX
DT 26-SEP-2001 (first entry)
XX
DE VEGF based monocyclic peptide 7.
XX
KW Human: VEGF, vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Disulfide-bond 1..9 /note="This bond cyclises the peptide"
FT XX
FT XX
PN WO200152875-A1.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01533.
XX
PR 18-JAN-2000; 2000US-0176293.
PR 16-MAY-2000; 2000US-0204550.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX

DR WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PI angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -

PS Claim 49; Page 32; 102pp; English.

CC The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the exposed loop of human
CC VEGFR (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidizing the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

SO Sequence 9 AA:

Query Match 100.0%; Score 50; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSVPPLTSSVC 9
DB 1 csvppltsvc 9

RESULT 2

ID AAB23826 standard; Protein; 1487 AA.

AC AAB23826;

DT 15-JAN-2001 (first entry)

DE Murine phospholipase A2 receptor protein SEQ ID NO:7.

KW Endocytic C lectin family; E-selectin; type C lectin; identification.

OS Mus sp.

PN US6117977-A.

PD 12-SEP-2000.

PF 24-APR-1997; 97US-0840062.

PR 24-APR-1996; 95US-0052524.

PR 24-APR-1996; 95US-0637021.

PA (GENENTECH) GENENTECH INC.

XX Wu K, Lasky LA;

DR WPI: 2000-601415/57.

XX New type C lectin protein having a signal sequence, cysteine-rich
PT domain, fibronectin type II domain, 8 type C lectin domains,
PT transmembrane domain and a short cytoplasmic domain, used as molecular
PT markers and hybridization probes -

PS Example; Fig 3; 72pp; English.

CC The present invention describes an isolated type C lectin polypeptide
CC (I) comprising amino acid residues 37-1393, 37-174, 175-229, 234-360,
CC 381-507, 520-645, 667-809, 824-951, 970-1108, 1110-1243, or 1259-1393
CC of the protein sequence given in AAB2382. The first 2 polypeptides are
CC capable of binding to a carbohydrate residue, and the rest of the
CC polypeptides are useful for producing antibodies capable of binding to
CC these 2 polypeptides. Polynucleotide sequence encoding the polypeptides
CC of the invention are useful in the identification and purification of
CC their native ligands, and as molecular markers of the tissues in which
CC they are expressed. They provide valuable sequence motifs, which can be
CC inserted or substituted into other native members of the endocytic type
CC C lectins, and provide hybridisation probes for searching cDNA and
CC genomic libraries for the coding sequence of other type C lectins.
CC Variants of type C lectins may be used therapeutically as competitive
CC inhibitors of the biological activity of native type C lectins.
CC The present sequence represents a murine phospholipase A2 receptor
CC protein which is homologous to a type C lectin.

SO Sequence 1487 AA:

Query Match 86.0%; Score 43; DB 21; Length 1487;
Best Local Similarity 77.8%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSVPPLTSSVC 9
DB 929 csvppltsvc 937

RESULT 3

ID AAB23826 standard; Protein; 73 AA.

AC AAB23826;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #24583.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIAXA CORP.

Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
L'Heuvelink J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.
DR N-PSDB: AAS59634.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 24882; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 73 AA;

Query Match 78.0%; Score 39; DB 22; Length 73;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPRTSVC 9
I I:||||
Db 35 csapltsv 43

RESULT 4
AAU53277 standard; Protein; 88 AA.
XX
AC AAU53277;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #14173.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI: 2001-616774/71.
DR N-PSDB: AAS59559.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 14472; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 88 AA;

Query Match 78.0%; Score 39; DB 22; Length 88;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSVPRTSVC 9
I I:|||| I
Db 43 csapltkpc 51

RESULT 5
AAU43894 standard; Protein; 101 AA.
XX
AC AAU43894;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #4790.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI

XX 09-AUG-2000; 2000WO-US21878.
PF
XX 09-AUG-1999; 99US-0147986.
PR 21-OCT-1999; 99US-0160807.
XX
XX (INCYTE GENOMICS INC.
PI Yang J;
PI Yue H, Lal P, Tang YF, Bandman O, Baughn MR, Azimzai Y, Lu DAM;
PI Yang J;
DR WPI: 2001-202760/20.
DR N-PSDB: AAF81739.
XX
XX New protease (inhibitors) useful for diagnosis and treatment of
PT autoimmune/inflammatory disorders such as acquired immunodeficiency
PT syndrome, Cushing's disease, Addison's disease and cell proliferative
PT disorders such as cancer -
XX
XX Claim 1; Page 114-115; 134pp; English.
PS
XX AAF81714 to AAF81740 encode the human proteases and protease inhibitors
CC (PRIMs) given in AAB74668 to AAB74694. The PRIMs can have activities such
CC as: anti-human immunodeficiency virus (HIV); antidiabetic; antithyroid;
CC immunostimulant; immunomodulator; antinflammatory; immunosuppressive;
CC nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;
CC fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;
CC virucide; antiparasitic; and hepatotropic. PRIM polynucleotide and
CC protein sequences can be used in the diagnosis, treatment and prevention
CC of autoimmune/inflammatory disorders such as AIDS, Digeorge's syndrome,
CC severe combined immunodeficiency disease (SCID), Chediak-Higashi
CC syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,
CC Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,
CC Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's
CC syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic
CC infections and cell proliferative disorder such as arteriosclerosis,
CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PRIM
CC polynucleotide sequences can be used in somatic or germline gene therapy
CC and in diagnosis of diseases. They can also be used in generating
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences and in molecular biology techniques.
XX
XX Sequence 742 AA:
SQ

Query Match 76.0%; Score 38; DB 22; Length 742;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPITVC 9
Db 726 csypitdpc 734
IIII:|
|
RESULT 8
AAE01782
ID AAE01782 standard; Protein; 802 AA.
XX
AC AAE01782;
XX
DT 17-JUL-2001 (first entry)
XX
DE Human gene 13 encoded secreted protein HDPNW93, SEQ ID NO:103.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder; gene therapy;
KW endocrine disorder; infection; wound healing; vulnery;

KW cell culture; chemotaxis; food additive; chromosome 17;
KW binding partner identification.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal_peptide
FT Protein 20..802
FT /label= Human_mature_secreted_protein
XX
XX WO200134627-A1.
XX
XX 17-MAY-2001.
XX
XX 08-NOV-2000; 2000WO-US30628.
XX
XX 12-NOV-1999; 99US-0164744.
XX 30-JUN-2000; 2000US-0215140.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Baker KP, Young PB;
XX WPI: 2001-316491/33.
XX N-PSDB: AAD05591.
XX
XX New nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX Parkinson's diseases and cancers -
XX
XX Claim 11; Page 492-494; 567pp; English.
PS
XX AAD05579-AAD05658 represent cDNAs corresponding to 28 human secreted
XX protein genes and AAE01770-AAE01849 represent the proteins they encode.
XX AAE01850-AAE01860 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 28 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention.
XX
XX Sequence 802 AA:
SQ

Query Match 76.0%; Score 38; DB 22; Length 802;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 CSVPITVC 9
Db 786 csypitdpc 794
IIII:|
|

RESULT 9
AAU58469
ID AAU58469 standard; Protein; 86 AA.
XX
AC AAU58469;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #19365.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neutroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.
XX
PA (COR-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
XX N-PSDB: AAS9591.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 19664; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 86 AA;

Query Match 74.0%; Score 37; DB 22; Length 86;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPLTSVC 9
||:|:| |

Db 39 csip1ghvc 47
RESULT 10
ABB67846
ID ABB67846 standard; Protein; 1929 AA.
XX
AC ABB67846;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 30330.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PERKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX N-PSDB: ABL11949.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 30330; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (AAB57737-AAB72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1929 AA;

Query Match 74.0%; Score 37; DB 22; Length 1929;
Best Local Similarity 55.6%; Pred. No. 8; 2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPLTSVC 9
||:|:| |

RESULT 11
AAG74036
ID AAG74036 standard; Protein; 90 AA.
XX
AC AAG74036;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4800.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX Homo sapiens.
XX WO200122920-A2.
XX PD 05-APR-2001.
XX PP 28-SEP-2000; 2000WO-US26524.
XX PR 29-SEP-1999; 99US-0157137.
XX PR 03-NOV-1999; 99US-0163280.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI: 2001-235357/24.
XX DR N-PSDB; AAH33467.
XX PS Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 11; Page 6589; 9803pp; English.
XX AAH32943 to AAH37195 and AAH37514 to AAH37788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAH37789 represent sequences used in the exemplification of the
XX present invention.
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX SO Sequence 90 AA:

Query Match 72.0%; Score 36; DB 22; Length 90;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPJTSVC 9
II:IIIIII
Db 46 ctfpilsvc 54

RESULT 12
ABBI0143
ID ABBI0143 standard; Protein; 133 AA.
XX
XX AC ABBI0143;
XX
XX 10-JAN-2002 (first entry)
XX
XX DE Human CDNA SEQ ID NO: 451.
XX
XX Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation.

XX Homo sapiens.
XX OS
XX PN WO200154474-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01349.
XX PR 31-JAN-2000; 2000US-179065P.
XX PR 04-FEB-2000; 2000US-180628P.
XX PR 24-FEB-2000; 2000US-184664P.
XX PR 02-MAR-2000; 2000US-186350P.
XX PR 16-MAR-2000; 2000US-189874P.
XX PR 17-MAR-2000; 2000US-190076P.
XX PR 18-APR-2000; 2000US-198123P.
XX PR 19-MAY-2000; 2000US-205515P.
XX PR 07-JUN-2000; 2000US-209467P.
XX PR 28-JUN-2000; 2000US-214866P.
XX PR 30-JUN-2000; 2000US-215135P.
XX PR 07-JUL-2000; 2000US-216647P.
XX PR 07-JUL-2000; 2000US-216880P.
XX PR 11-JUL-2000; 2000US-217487P.
XX PR 11-JUL-2000; 2000US-217496P.
XX PR 14-JUL-2000; 2000US-218290P.
XX PR 26-JUL-2000; 2000US-220963P.
XX PR 14-AUG-2000; 2000US-220964P.
XX PR 14-AUG-2000; 2000US-224518P.
XX PR 14-AUG-2000; 2000US-224519P.
XX PR 14-AUG-2000; 2000US-225213P.
XX PR 14-AUG-2000; 2000US-225214P.
XX PR 14-AUG-2000; 2000US-225266P.
XX PR 14-AUG-2000; 2000US-225267P.
XX PR 14-AUG-2000; 2000US-225268P.
XX PR 14-AUG-2000; 2000US-225270P.
XX PR 14-AUG-2000; 2000US-225477P.
XX PR 14-AUG-2000; 2000US-225757P.
XX PR 14-AUG-2000; 2000US-225758P.
XX PR 14-AUG-2000; 2000US-225759P.
XX PR 18-AUG-2000; 2000US-226279P.
XX PR 22-AUG-2000; 2000US-226681P.
XX PR 22-AUG-2000; 2000US-226868P.
XX PR 22-AUG-2000; 2000US-227182P.
XX PR 23-AUG-2000; 2000US-227009P.
XX PR 30-AUG-2000; 2000US-228924P.
XX PR 01-SEP-2000; 2000US-229287P.
XX PR 01-SEP-2000; 2000US-229343P.
XX PR 01-SEP-2000; 2000US-229344P.
XX PR 01-SEP-2000; 2000US-229345P.
XX PR 05-SEP-2000; 2000US-229509P.
XX PR 05-SEP-2000; 2000US-229513P.
XX PR 06-SEP-2000; 2000US-230437P.
XX PR 06-SEP-2000; 2000US-230438P.
XX PR 08-SEP-2000; 2000US-231242P.
XX PR 08-SEP-2000; 2000US-231243P.
XX PR 08-SEP-2000; 2000US-231244P.
XX PR 08-SEP-2000; 2000US-231413P.
XX PR 08-SEP-2000; 2000US-231414P.
XX PR 08-SEP-2000; 2000US-232080P.
XX PR 08-SEP-2000; 2000US-232081P.
XX PR 12-SEP-2000; 2000US-231968P.
XX PR 14-SEP-2000; 2000US-232397P.
XX PR 14-SEP-2000; 2000US-232398P.
XX PR 14-SEP-2000; 2000US-232399P.
XX PR 14-SEP-2000; 2000US-232400P.
XX PR 14-SEP-2000; 2000US-232401P.
XX PR 14-SEP-2000; 2000US-233063P.
XX PR 14-SEP-2000; 2000US-233064P.
XX PR 14-SEP-2000; 2000US-233065P.
XX PR 21-SEP-2000; 2000US-234223P.
XX PR 21-SEP-2000; 2000US-234274P.
XX PR 25-SEP-2000; 2000US-234997P.
XX PR 25-SEP-2000; 2000US-234998P.

PR 26-SEP-2000; 2000US-235484P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 27-SEP-2000; 2000US-235836P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 13-OCT-2000; 2000US-239937P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241221P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241786P.
 PR 20-OCT-2000; 2000US-241787P.
 PR 20-OCT-2000; 2000US-241808P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 20-OCT-2000; 2000US-241826P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 08-NOV-2000; 2000US-246474P.
 PR 08-NOV-2000; 2000US-246475P.
 PR 08-NOV-2000; 2000US-246476P.
 PR 08-NOV-2000; 2000US-246477P.
 PR 08-NOV-2000; 2000US-246478P.
 PR 08-NOV-2000; 2000US-246523P.
 PR 08-NOV-2000; 2000US-246524P.
 PR 08-NOV-2000; 2000US-246525P.
 PR 08-NOV-2000; 2000US-246526P.
 PR 08-NOV-2000; 2000US-246527P.
 PR 08-NOV-2000; 2000US-246528P.
 PR 08-NOV-2000; 2000US-246532P.
 PR 08-NOV-2000; 2000US-246609P.
 PR 08-NOV-2000; 2000US-246610P.
 PR 08-NOV-2000; 2000US-246611P.
 PR 17-NOV-2000; 2000US-249207P.
 PR 17-NOV-2000; 2000US-249208P.
 PR 17-NOV-2000; 2000US-249209P.
 PR 17-NOV-2000; 2000US-249210P.
 PR 17-NOV-2000; 2000US-249211P.
 PR 17-NOV-2000; 2000US-249212P.
 PR 17-NOV-2000; 2000US-249213P.
 PR 17-NOV-2000; 2000US-249214P.
 PR 17-NOV-2000; 2000US-249215P.
 PR 17-NOV-2000; 2000US-249216P.
 PR 17-NOV-2000; 2000US-249217P.
 PR 17-NOV-2000; 2000US-249218P.
 PR 17-NOV-2000; 2000US-249244P.
 PR 17-NOV-2000; 2000US-249245P.
 PR 17-NOV-2000; 2000US-249246P.
 PR 17-NOV-2000; 2000US-249255P.
 PR 17-NOV-2000; 2000US-249256P.
 PR 17-NOV-2000; 2000US-249257P.
 PR 17-NOV-2000; 2000US-249259P.
 PR 01-DEC-2000; 2000US-250160P.
 PR 01-DEC-2000; 2000US-250391P.
 PR 05-DEC-2000; 2000US-251030P.
 PR 05-DEC-2000; 2000US-251988P.
 PR 05-DEC-2000; 2000US-256719P.
 PR 06-DEC-2000; 2000US-251479P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 PR 08-DEC-2000; 2000US-251989P.
 PR 08-DEC-2000; 2000US-251990P.
 PR 11-DEC-2000; 2000US-254097P.
 PR 05-JAN-2001; 2001US-259678P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM.
 XX WPI: 2001-476161/51.
 DR N-PSDB: ABA06355.
 XX
 PT Isolated nucleic acid molecule encoding an inflammation-associated
 PT polypeptide is used in preventing, treating or ameliorating a medical
 PT condition _
 XX
 PS Claim 11; SEQ ID NO: 451; 859pp + Sequence Listing; English.
 XX
 CC The present invention provides human cDNAs, proteins and related genomic
 CC DNAs. These can be used in the treatment of neural, immune system,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders and inflammation. The present sequence
 CC is a protein of the invention.
 XX
 SQ Sequence 133 AA:

Query Match 72.0%; Score 36; DB 22; Length 133;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPPLTVC 9
 |||| :||
 DB 74 csvpplvc 82

RESULT 13
 AAU18408
 ID AAU18408 standard; Protein; 133 AA.

AC AAU18408;

DT 21-NOV-2001 (first entry)

DE Human endocrine polypeptide SEQ ID No 363.

XX
 KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
 KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
 KW cerebroprotective; nootropic; antibacterial; vitucide; fungicide; cancer;
 KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
 KW cerebrovascular disorder; nervous system disorder; bacterial infection;
 KW fungal infection; viral infection; ocular disorder; endocrine disorder;
 KW gastrointestinal disorder; renal disorder; respiratory disorder;
 KW wound healing; skin aging; organ transplantation; food preservative;
 KW tissue regeneration; anti-infertility.

OS Homo sapiens.

PN WO200155364-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01308.

PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232402.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234597.
 PR 25-SEP-2000; 2000US-0234598.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-451936/48.
 N-PSDB; AAS29637.
 Isolated polypeptide for treating, preventing and/or prognosing
 disorders of the endocrine system such as reproductive disorders,
 endocrine cancers and also for testing and detection e.g. diagnosis -
 Claim 11; SEQ ID NO 363; 604pe; English.
 Sequences AAU18282-AAU18507 represent endocrine polypeptides of the
 invention. Endocrine polypeptides and their associated polynucleotides
 are useful in the diagnosis, treatment and prevention of various types of
 disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 chickens or sheep. A pathological condition can be determined by
 determining the presence or absence of a mutation in an endocrine
 polynucleotide. The treatable disorders include autoimmune diseases such
 as rheumatoid arthritis, hyperproliferative disorders such as neoplasms

CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 72.0% Score 36; DB 22; Length 133;
Best Local Similarity 66.7% Pred. NO. 88;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSVPLTFSVC 9
||| :||
Db 74 cswpsanvc 82

RESULT 14

AAM43498
ID AAM43498 standard; Protein; 133 AA.

XX AAM43498:

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 176.

XX Human: antiarthritic; antirheumatic; antiproliferative; vasotropic;
KM cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KM fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KM neuroprotective; anti-allergic; hepatotropic; antidiabetic;
KM antiinflammatory; anticancer; vulnary; anticonvulsant; antibacterial;
KM antiparasitic; cardiac; gene therapy; cancer; immune disorder;
KM cardiovascular disorder; neurological disease; infection; human.

XX Homo sapiens.

XX WO200155308-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01309.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234273.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240360.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-488781/53.
DR N-PSDB: AA163804.
XX
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
XX
XX
XX
XX
PS Claim 11; SEQ ID NO 176; 664pp + Sequence Listing; English.
CC
CC The invention relates to human polynucleotides (AA163803-AA164012) and
CC the encoded proteins (AA163803-AA164012) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pt_sequences.
XX
XX
SQ Sequence 133 AA;

Query Match 72.0%; Score 36; DB 22; Length 133;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSVPPTSVC 9
DB 74 csypanvc 82
RESULT 15
AAU19674
ID AAU19674 standard; Protein; 133 AA.
XX
AC AAU19674;
XX
DT 06-DEC-2001 (first entry)
DE
DE Human novel extracellular matrix protein, Seq ID No 324.
XX
XX Human; secreted extracellular matrix protein; immunomodulatory;
XX Anti-HIV; antineoplastic; antineoplastic; cardiact; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; cytoprotective;
XX antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
XX human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
XX cancers; hyperproliferative disorder; breast neoplasm; melanoma;
XX Sezary syndrome; Gaucher's disease; neurological diseases;
XX Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
XX cardiac arrest; tachycardia; angina; infection; corneal infections;
XX wound healing; immunogen; gene therapy; antisense; food additive.
XX
XX Homo sapiens.
XX
XX W0200155368-A1.
XX
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01348.
XX
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0231480.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233065.
 PR 14-SEP-2000; 2000US-0233066.
 PR 21-SEP-2000; 2000US-0234222.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249399.
 PR 17-NOV-2000; 2000US-0249500.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI. 2001-465572/50.
 DR N-PSDB; MAS31245.
 XX
 XX
 PT Nucleic acid molecules encoding human secreted extracellular matrix
 PR proteins, used in preventing, treating or ameliorating a disorder, e.g.
 PT Alzheimer's and Parkinson's diseases and cancers -
 XX
 XX
 PS Claim 11; SEQ ID NO 324; 577pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules encoding
 CC novel human secreted extracellular matrix proteins (SPs). The
 CC polynucleotides and proteins are used to prevent, treat a medical
 CC condition in e.g. humans, mice, rabbits, goats, cats, dogs,
 CC chickens or sheep. For example, disorders associated with decreased
 CC expression of SPs. The SP polynucleotide or a vector expressing them may
 CC be administered to treat diseases by gene therapy. Antisense molecules
 CC may be administered to down regulate expression of SPs by binding with
 CC the cells own genes and preventing their expression. The polynucleotides
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be
 CC used as antigens to produce antibodies and to identify modulators
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses
 CC include wound healing, maintenance of organs before transplantation,
 CC

Query Match 72.0%; Score 36; DB 22; Length 133;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CSVPLTSCV 9

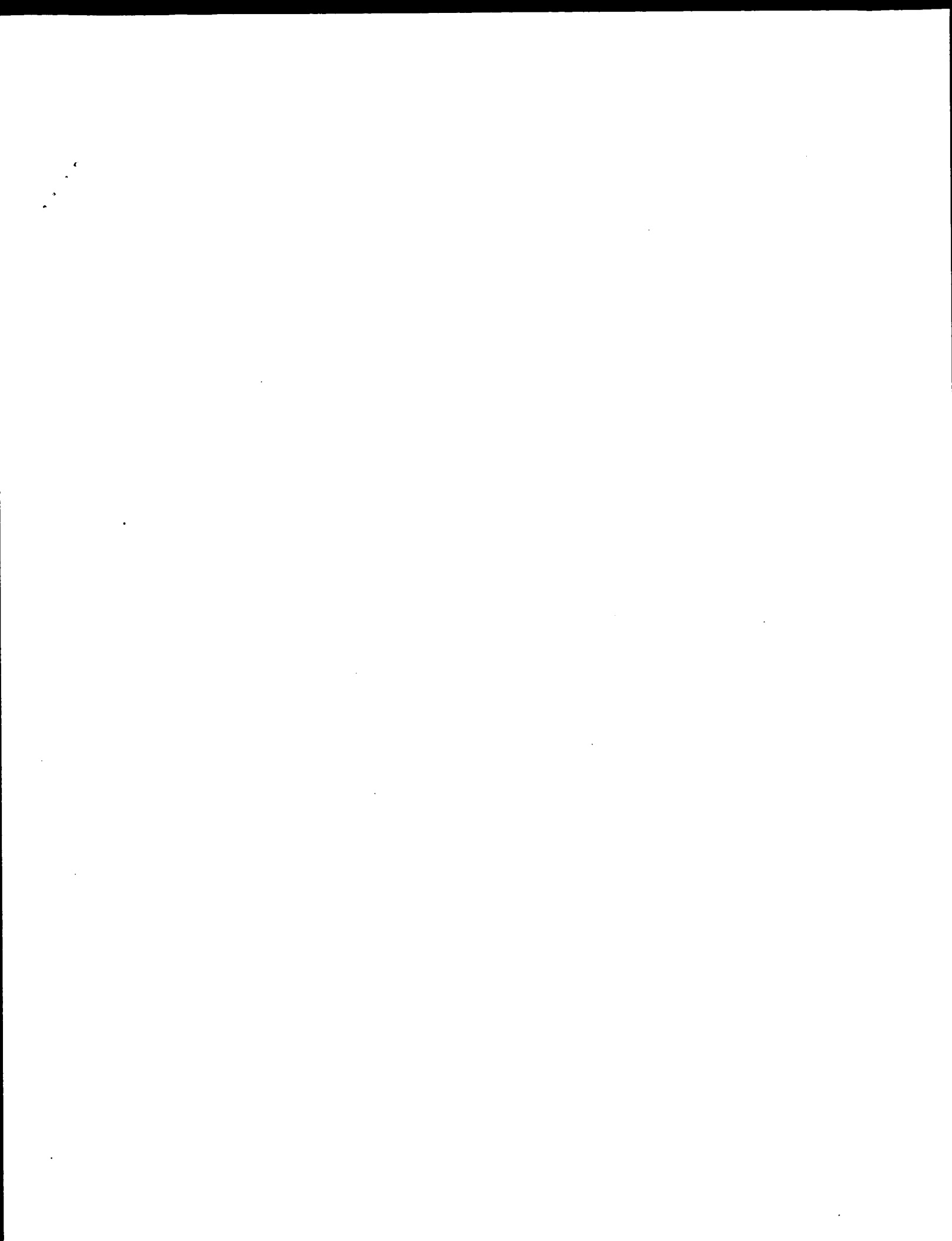
Mon Jun 17 16:10:51 2002

us-09-761-636a-10.open.rag

Page 13

Db 74 csypsanhc 82

Search completed: June 17, 2002, 16:02:14
Job time: 423 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:37 ; Search time 75.98 seconds
(without alignments)
2.893 Million cell updates/sec

Title: US-09-761-636A-10
Perfect score: 50
Sequence: 1 CSVPPLTSSVC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	86.0	1487	3	US-08-840-062-7
2	36	72.0	201	2	US-08-688-342-1
3	36	72.0	201	2	US-09-113-788-1
4	35	70.0	359	3	US-08-586-165-4
5	34	68.0	386	1	US-08-134-012-3
6	34	68.0	386	1	US-08-520-519-3
7	34	68.0	747	2	US-08-616-693A-51
8	34	68.0	747	3	US-08-885-291-51
9	34	68.0	747	4	US-09-496-672-51
10	32	64.0	134	4	US-08-543-246B-20
11	32	64.0	216	4	US-08-543-246B-9
12	32	64.0	216	4	US-08-543-246B-24
13	32	64.0	321	4	US-08-915-795-9
14	32	64.0	325	4	US-08-915-795-3
15	32	64.0	354	4	US-08-915-795-5
16	32	64.0	358	4	US-08-915-795-8
17	32	64.0	379	2	US-07-857-224B-87
18	32	64.0	505	1	US-08-220-603A-10
19	32	64.0	1463	1	US-08-220-603A-11
20	31	62.0	14	1	US-08-297-633A-5
21	31	62.0	14	2	US-08-485-721-5
22	31	62.0	14	2	US-08-392-935-5
23	31	62.0	14	3	US-08-897-236-5
24	31	62.0	14	4	US-09-167-874-5
25	31	62.0	14	5	PCT-US93-08325-5
26	31	62.0	14	5	PCT-US93-08326-5
27	31	62.0	80	5	PCT-US95-06266-130

28	31	62.0	85	2	US-08-480-229C-7	Sequence 7, Appl
29	31	62.0	85	2	US-08-659-235C-7	Sequence 7, Appl
30	31	62.0	87	2	US-08-485-721-11	Sequence 11, Appl
31	31	62.0	87	2	US-08-392-935-11	Sequence 11, Appl
32	31	62.0	87	5	PCT-US93-08326-11	Sequence 11, Appl
33	31	62.0	92	2	US-08-737-925-7	Sequence 26, Appl
34	31	62.0	104	1	US-08-111-939-26	Sequence 27, Appl
35	31	62.0	108	2	US-08-162-402B-27	Sequence 27, Appl
36	31	62.0	222	2	US-08-485-721-9	Sequence 9, Appl
37	31	62.0	222	2	US-08-392-935-9	Sequence 9, Appl
38	31	62.0	222	5	PCT-US93-08326-9	Sequence 9, Appl
39	31	62.0	232	2	US-08-485-721-2	Sequence 2, Appl
40	31	62.0	232	2	US-08-392-935-2	Sequence 2, Appl
41	31	62.0	232	3	US-08-897-236-2	Sequence 2, Appl
42	31	62.0	232	3	US-08-897-236-11	Sequence 11, Appl
43	31	62.0	232	4	US-09-167-874-2	Sequence 2, Appl
44	31	62.0	232	4	US-09-167-874-11	Sequence 11, Appl
45	31	62.0	232	5	PCT-US93-08326-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-840-062-7
; Sequence 7, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: WU, KAI
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840, 062
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1487 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-840-062-7

Query Match 86.0%; Score 43; DB 3; Length 1487;

Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSVPPLTSSVC 9
|||||1:1
Db 929 CSVPPLTSSVC 937

RESULT 2
US-08-688-342-1
; Sequence 1, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: MMLR1DT01
; CLONE: 515847
; US-08-688-342-1

Query Match 72.0%; Score 36; DB 2; Length 201;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 C5VPLT5VC 9
Db 187 C5VPSYSIC 195

RESULT 3
US-09-113-788-1
; Sequence 1, Application US/09113788
; Patent No. 5969104
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,788
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/688,342
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: MMLR1DT01
; CLONE: 515847
; US-09-113-788-1

Query Match 72.0%; Score 36; DB 2; Length 201;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 C5VPLT5VC 9
Db 187 C5VPSYSIC 195

RESULT 4
US-08-586-165-4
; Sequence 4, Application US/08586165
; Patent No. 6054298
; GENERAL INFORMATION:
; APPLICANT: Laufer, Edward M.
; APPLICANT: Orozco, Olivia E.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Fringe Proteins and Pattern Formation
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,165
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540

ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,693A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5874241thrup, Thomas E
REGISTRATION NUMBER: 33,268
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-816-693A-51

Query Match 68.0%; Score 34; DB 2; Length 747;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPPLTSCV 9
Db 215 CRVPLGKVC 223

RESULT 8
US-08-885-291-51
Sequence 51, Application US/0885291A
Patent No. 6057125
GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S.
APPLICANT: Turek, Fred W.
APPLICANT: Pinto, Lawrence H.
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
FILE REFERENCE: 0290-5
CURRENT APPLICATION NUMBER: US/08/885,291A
CURRENT FILING DATE: 1997-06-30
EARLIER APPLICATION NUMBER: 08/816,693
EARLIER FILING DATE: 1997-03-13
NUMBER OF SEQ ID NOS: 55
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 51
LENGTH: 747
TYPE: PRT
ORGANISM: Mus musculus
US-08-885-291-51

Query Match 68.0%; Score 34; DB 3; Length 747;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPPLTSCV 9
Db 215 CRVPLGKVC 223

RESULT 9
US-09-496-672-51
Sequence 51, Application US/09496672
Patent No. 6291429
GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S.

APPLICANT: Turek, Fred W.
APPLICANT: Pinto, Lawrence H.
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
FILE REFERENCE: 0290-5
CURRENT APPLICATION NUMBER: US/09/496,672
CURRENT FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: 08/885,291
PRIOR FILING DATE: 1997-06-30
PRIOR APPLICATION NUMBER: 08/816,693
PRIOR FILING DATE: 1997-03-13
NUMBER OF SEQ ID NOS: 55
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 51
LENGTH: 747
TYPE: PRT
ORGANISM: Mus musculus
US-09-496-672-51

Query Match 68.0%; Score 34; DB 4; Length 747;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSVPPLTSCV 9
Db 215 CRVPLGKVC 223

RESULT 10
US-08-543-246B-20
Sequence 20, Application US/08543246B
Patent No. 6262244
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: DNA and amino acid sequence specific for
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michael W. Glynn
ADDRESS: No. 6262244artis Corporation
STREET: 564 Morris Avenue
CITY: Summit,
STATE: NJ
COUNTRY: US
ZIP: 07901-1027
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/543,246B
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,663
FILING DATE: 28-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02469
FILING DATE: 27-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,514
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kassenooff, Melvyn M.
REGISTRATION NUMBER: 26,389
REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-522-6927
TELEFAX: 908-522-6955
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 134 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: C-terminal
US-08-543-246B-20

Query Match 64.0%; Score 32; DB 4; Length 134;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSVP1TSVC 9
DB 121 CSTPNTYIC 129

RESULT 11
US-08-543-246B-9
Sequence 9, Application US/08543246B
Patent No. 6262244
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA and amino acid sequence specific for
TITLE OF INVENTION: natural killer cells
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michael W. Glynn
ADDRESSEE: No. 6262244artis Corporation
STREET: 564 Morris Avenue
CITY: Summit,
STATE: NJ
COUNTRY: US
ZIP: 07901-1027
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/543,246B
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,663
FILING DATE: 28-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02469
FILING DATE: 27-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,514
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kassenoit, Melvyn M.
REGISTRATION NUMBER: 26,389
REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-522-6927
TELEFAX: 908-522-6955
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 216 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-543-246B-9

Query Match 64.0%; Score 32; DB 4; Length 216;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSVP1TSVC 9

DB 203 CSTPNTYIC 211

RESULT 12
US-08-543-246B-24
Sequence 24, Application US/08543246B
Patent No. 6262244
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA and amino acid sequence specific for
TITLE OF INVENTION: natural killer cells
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michael W. Glynn
ADDRESSEE: No. 6262244artis Corporation
STREET: 564 Morris Avenue
CITY: Summit,
STATE: NJ
COUNTRY: US
ZIP: 07901-1027
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/543,246B
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/676,663
FILING DATE: 28-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02469
FILING DATE: 27-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,514
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kassenoit, Melvyn M.
REGISTRATION NUMBER: 26,389
REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-522-6927
TELEFAX: 908-522-6955
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 216 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: C-terminal
US-08-543-246B-24

Query Match 64.0%; Score 32; DB 4; Length 216;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSVP1TSVC 9
DB 203 CSTPNTYIC 211

RESULT 13
US-08-915-795-9
Sequence 9, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN

APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-9

Query Match 64.0%; Score 32; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSV 8
Db 169 SVPLTSV 175

RESULT 14
US-08-915-795-3
Sequence 3, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
TISSUE TYPE: Human Breast
US-08-915-795-3

Query Match 64.0%; Score 32; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVPLTSV 8
Db 140 SVPLTSV 146

RESULT 15
US-08-915-795-5
Sequence 5, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
TYPE: amino acid

```

;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEITICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Lung
;
US-08-915-795-5

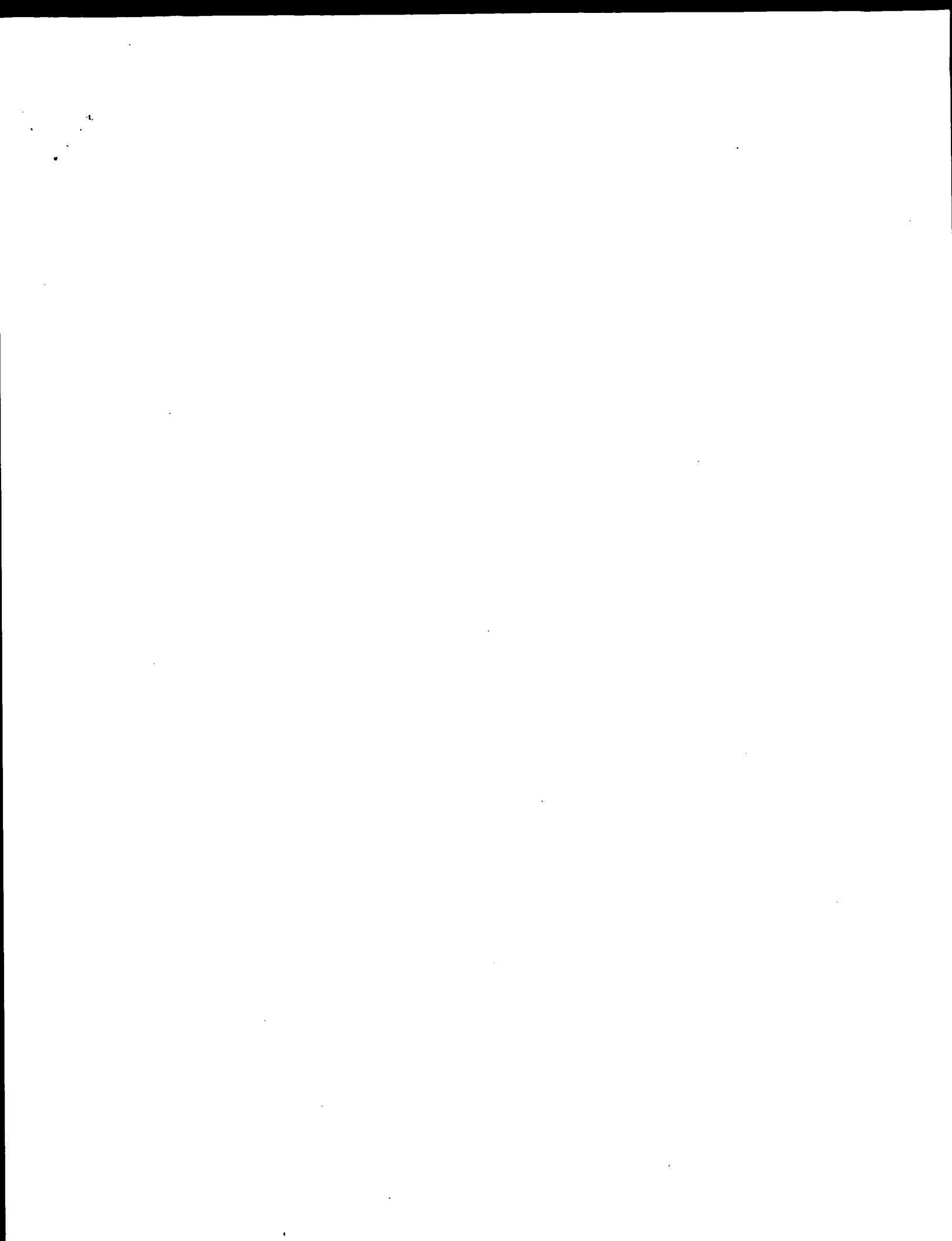
```

```

Query Match          64.0%; Score 32; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 SVPLTSV 8
    |||||
Db 169 SVPLTSV 175

```

Search completed: June 17, 2002, 15:56:38
 Job time: 87 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:24 ; Search time 95.43 Seconds
(without alignments)
7.048 Million cell updates/sec

Title: us-09-761-636a-11
Perfect score: 42
Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0'
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : PIR.71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	92.9	2543	2 T31687	surface antigen - p
2	38	90.5	380	2 D44490	retrovirus-related
3	36	85.7	980	2 T00045	cellodextrin phosp
4	36	85.7	2533	2 T28675	alpha-51D immobili
5	36	85.7	2533	2 T28674	alpha-51D-immobili
6	34	81.0	1506	2 T30886	integumentary muc
7	34	81.0	1513	2 A54895	mucin 2, intestina
8	34	81.0	2476	2 T34022	zonadhesin - pig
9	33	78.6	55	2 S66336	protein kinase AK2
10	33	78.6	191	2 I46412	keratin KAP5.4 - s
11	33	78.6	207	2 A64655	hypothetical prote
12	33	78.6	207	2 B71943	hypothetical prote
13	33	78.6	223	2 B38346	ultra-high-sulfur
14	33	78.6	230	2 A38346	ultra-high-sulfur
15	33	78.6	303	2 WZBEM6	gene 20 protein -
16	33	78.6	303	2 C33374	hypothetical prote
17	33	78.6	515	2 T23089	hypothetical prote
18	33	78.6	647	2 T23814	hypothetical prote
19	33	78.6	670	2 F88297	protein M28.1 (imp
20	33	78.6	891	2 H86306	F20D23.20 protein
21	33	78.6	1227	2 T49963	hypothetical prote
22	33	78.6	3020	2 A43932	mucin 2 precursor,
23	33	78.6	3133	2 S52093	hemocytin - silkw
24	32	76.2	77	2 JC4790	protease inhibitor
25	32	76.2	270	2 T47421	hypothetical prote
26	32	76.2	303	2 T42933	hypothetical prote
27	32	76.2	305	2 B53782	peroxisome assembl
28	32	76.2	452	2 F84712	hypothetical prote
29	32	76.2	491	2 A49179	melanoma antigen h

30	32	76.2	652	2 I38400	melanoma-associate
31	32	76.2	139	2 A41234	melanocyte-specific
32	32	76.2	1319	2 S55598	tegument protein 0
33	32	76.2	2044	2 T13704	still life protein
34	32	76.2	2180	2 T29764	hypothetical prote
35	32	76.2	2395	1 S50870	surface protein ty
36	32	76.2	5376	2 T42215	zonadhesin - mouse
37	31	73.8	63	2 S08572	chymotrypsin/elast
38	31	73.8	110	2 S16496	hypothetical prote
39	31	73.8	140	2 A33787	vascular endotheli
40	31	73.8	128	2 I51295	vascular endotheli
41	31	73.8	146	2 S57956	ovine vascular end
42	31	73.8	161	2 T27849	hypothetical prote
43	31	73.8	169	1 S18946	ultra-high-sulfur
44	31	73.8	182	2 A36866	ultra-high-sulfur
45	31	73.8	190	2 S52130	vascular endotheli

ALIGNMENTS

RESULT 1
T31687
surface antigen - Parametium primaurelia
C:Species: Parametium primaurelia
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001
C:Accession: T31687
R:Bourgain-Guilliet, F.; Caron, F.
Journal of Eukaryot. Microbiol. 43, 303-314, 1996
A>Title: Molecular characterization of the D surface protein gene subfamily in Par
A:Reference number: Z21061; MUID:96313351
A:Accession: T31687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2543 <BOU>
A:Cross-references: EMBL:X96616; NID:91235576; PIDN:CA65436.1
C:Genetics:
A:Genetic code: SGC5
A:Superfamily: G surface protein

Query Match 92.9% Score 39; DB 2; Length 2543;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1 CVPLTSC 7
DB	2158 CIPRTSC 2164

RESULT 2
D44490
retrovirus-related reverse transcriptase homolog (clone NVC) - pteromalid wasp (Na
C:Species: Nasonia vitripennis
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: D44490; F40442
R:Burke, W.D.; Eickbush, D.G.; Xiong, Y.; Jakubczak, J.; Eickbush, T.H.
Mol. Biol. Evol. 10, 163-185, 1993
A>Title: Sequence relationship of retrotransposable elements R1 and R2 within and
A:Reference number: A44490; MUID:91196484
A:Contents: retrotransposable element R1
A:Accession: D44490
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-380 <BUR>
A>Note: sequence extracted from NCBI backbone (NCBI:127237)
R:Jakubczak, J.L.; Burke, W.D.; Eickbush, T.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 3295-3299, 1991
A>Title: Retrotransposable elements R1 and R2 interrupt the rRNA genes of most ins
A:Reference number: A40442; MUID:91195337
A:Accession: F40442
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptu
A:Molecule type: DNA

A:Residues: 304-320 <JAK>
C:Superfamily: silkworm pol protein

Query Match 90.5%; Score 38; DB 2; Length 380;
Best Local Similarity 85.7%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPITSC 7
|:|:|:|
DB 356 CVPITSC 362

RESULT 3
T00045
cellohextrin phosphorylase - Clostridium thermocellum
C:Species: Clostridium thermocellum
C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 15-Oct-1999

R:Kawaguchi, T.; Ikeuchi, Y.; Tsutsumi, N.; Kan, A.; Sumitani, J.; Arai, M.
J. Ferment. Bioeng. 85, 144-149, 1998
A:Title: Cloning, nucleotide sequence, and expression of the Clostridium thermocellum ce
A:Reference number: 214077
A:Accession: T00045
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-980 <KAW>
A:Cross-references: EMBL:AB006822; NID:d1117395; PIDN:BA42081.1; PID:d1022940
A:Experimental source: ATCC 27405
C:Genetics:

A:Gene: cdp

Query Match 85.7%; Score 36; DB 2; Length 980;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPITSC 7
|:|:|:|
DB 550 CIPMTAC 556

RESULT 4
T28675
alpha-51D immobilization antigen - Paramecium tetraurelia
C:Species: Paramecium tetraurelia
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T28675
R:Schwegmann, K.J.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20506
A:Accession: T28675
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-2533 <SCH>
A:Cross-references: EMBL:X96400; PIDN:CAA65264.1
C:Genetics:

A:Gene: alpha-51D
A:Genetic code: SGC5
A:Introns: 280/3; 538/2; 1248/2
C:Superfamily: G surface protein

Query Match 85.7%; Score 36; DB 2; Length 2533;
Best Local Similarity 57.1%; Pred. No. 98;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPITSC 7
|:|:|:|
DB 2149 CIPITNC 2155

RESULT 5

T28674

alpha-51D-immobilization antigen - Paramecium tetraurelia

C:Species: Paramecium tetraurelia

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T28674

R:Schmidt, H.J.

submitted to the EMBL Data Library, March 1995

A:Reference number: Z20505

A:Accession: T28674

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2533 <SCH>

A:Cross-references: EMBL:X85135; NID:g728634; PID:g728635; PIDN:CAA59447.1

C:Genetics:

A:Genetic code: SGC5

A>Note: alpha-51D

C:Superfamily: G surface protein

Query Match 85.7%; Score 36; DB 2; Length 2533;
Best Local Similarity 57.1%; Pred. No. 98;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPITSC 7
|:|:|:|
DB 2149 CIPITNC 2155

RESULT 6

T30886
Integumentary mucin B.1 - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 03-Nov-2000

C:Accession: T30886

R:Juba, W.; Hoffmann, W.

J. Biol. Chem. 272, 1805-1810, 1997

A:Title: Similarities of Integumentary mucin B.1 (FIM-B.1) from Xenopus laevis and

A:Reference number: Z20920; M0ID:97153143

A:Accession: T30886

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1506 <JOB>

A:Cross-references: EMBL:Y08296; NID:g1839051; PIDN:CAA69604.1; PID:g1839052

C:Genetics:

A>Note: FIM-B.1

C:Superfamily: pig submaxillary mucin

Query Match 81.0%; Score 34; DB 2; Length 1506;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPITSC 7
|:|:|:|
DB 1292 CVPITSC 1298

RESULT 7

A54895
mucin 2, intestinal/tracheal - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000

C:Accession: A54895

R:Ohnori, H.; Dohman, A.F.; Gallup, M.; Tsuda, T.; Kai, H.; Gum Jr., J.R.; Klm, Y.

J. Biol. Chem. 269, 17833-17840, 1994

A:Title: Molecular cloning of the amino-terminal region of a rat MUC 2 mucin gene h

A:Reference number: A54895; M0ID:94299489

A:Accession: A54895

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1513 <OHM>

A:Cross-references: GB:U07615

C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology;

C:Keywords: intestine

Query Match 81.0% Score 34; DB 2; Length 1513;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 |||||
 Db 709 CVPITSC 715

RESULT 8
 T34022
 zonadhesin - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T34022
 R:Hardy, D.M.; Garbers, D.L.
 J. Biol. Chem. 270, 26025-26028, 1995
 A:Title: A sperm membrane protein that binds in a species-specific manner to the egg
 A:Reference number: Z21464; MUID:96064658
 A:Accession: T34022
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2476 <HAR>
 A:Cross-references: EMBL:U40024; NID:g1066465; PID:g1066466; PIDN:AAC4846.1
 A:Experimental source: strain Meishan; testis
 C:Genetics:
 A:Gene: Zan
 A:Function:
 A:Description: may be involved in sperm adhesion to the zona pellucida

Query Match 81.0% Score 34; DB 2; Length 2476;
 Best Local Similarity 71.4%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 |||||
 Db 1505 CVPITSC 1511

RESULT 9
 S66336
 protein kinase AK23 (EC 2.7.1.-) - Arabidopsis thaliana (fragment)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 24-Sep-1999
 C:Accession: S66336; S58262
 R:Thummler, F.; Kirchner, M.; Teuber, R.; Dittlich, P.
 Plant Mol. Biol. 29, 551-565, 1995
 A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes
 A:Reference number: S66314; MUID:96123233
 A:Accession: S66336
 A:Molecule type: DNA
 A:Residues: 1-55 <THD>
 A:Cross-references: EMBL:X86968; NID:g928913; PIDN:CAA60531.1; PID:g928914
 C:Genetics:
 A:Gene: AK23
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
 C:Keywords: ATP; phosphotransferase; protein kinase
 F:1-55/Domain: protein kinase homology (fragment) <KIN>

Query Match 78.6% Score 33; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITSC 6
 |||||
 Db 27 CVPITSC 32

RESULT 10

I46412

keratin KAP5.4 - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 24-Sep-1999

C:Accession: I46412; S34215

R:Jenkins, B.J.; Powell, B.C.

J. Invest. Dermatol. 103, 310-317, 1994

A:Title: Differential expression of genes encoding a cysteine-rich keratin family

A:Reference number: I46412; MUID:94336466

A:Accession: I46412

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-191 <JEN>

A:Cross-references: EMBL:X73434; NID:g313719; PIDN:CAA51829.1; PID:g313720

C:Genetics:

A:Gene: KRTAP5.4

C:Superfamily: ultra-high-sulfur keratin

Query Match 78.6% Score 33; DB 2; Length 191;
 Best Local Similarity 71.4%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 |||||
 Db 91 CVPITSC 97

RESULT 11

A64655

hypothetical protein HP1081 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: A64655

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; Mc

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wether

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpi, P.D.; Smith, H.O.; Frase

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: A64655

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-207 <TOM>

A:Cross-references: GB:AE000615; GB:AE000511; NID:g2314230; PIDN:AAD08133.1; PID:g2314230

Query Match

78.6% Score 33; DB 2; Length 207;

Best Local Similarity 71.4%; Pred. No. 43;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 |||||
 Db 134 CVPITSC 140

RESULT 12

B71943

hypothetical protein jhp0344 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999

C:Accession: B71943

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith,

Ives, C.; Gibson, R.; Weiberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyis, G

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric

A:Reference number: A71800; MUID:99120557

A:Accession: B71943

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-207 <ARN>
A:Cross-references: GB:AE001470; GB:AE001439; NID:g4154869; PIDN:AAD05923.1; PID:g415487
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0344

Query Match 78.6%; Score 33; DB 2; Length 207;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPITSC 7
111:11
DB 134 CVPOTAC 140

RESULT 13
B38346
ultra-high-sulfur keratin 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: A38660; B38346
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogelt, G.
J. Biol. Chem. 266, 4024, 1991
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38660; MUID:91154184
A:Accession: A38660
A:Molecule type: DNA
A:Residues: 1-223 <MO2>
A:Cross-references: GB:M37760; NID:g200963; PIDN:AAA40107.1; PID:g200964
A:Note: This is a correction
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogelt, G.
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38346; MUID:91065960
A:Accession: B38346
A:Molecule type: DNA
A:Residues: 1-21, 'GGCGSGCGCGCGCGSSCKPVCC', 22-40, 'GSS', 44-45, 'G', 47-48, 'S', 50, 'GSS',
<MOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
A:Note: the sequence reported in this paper has been corrected. See A38660
C:Superfamily: ultra-high-sulfur keratin

Query Match 78.6%; Score 33; DB 2; Length 223;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPITSC 7
111:11
DB 28 CVPVCSC 34

RESULT 14
A38346
ultra-high-sulfur keratin 1 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 24-Sep-1999
C:Accession: A38346
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogelt, G.
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin
A:Reference number: A38346; MUID:91065960
A:Accession: A38346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <MOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
C:Superfamily: ultra-high-sulfur keratin

Query Match 78.6%; Score 33; DB 2; Length 230;
Best Local Similarity 71.4%; Pred. No. 47;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CVPITSC 7
111:11
DB 55 CVPVCSC 61

RESULT 15
W2BEM6
gene 20 protein - saimiriine herpesvirus 1 (strain 11)
C:Species: saimiriine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C:Accession: B36808
R:Albrecht, J.
submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36808
A:Accession: B36808
A:Molecule type: DNA
A:Residues: 1-303 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45644.1; PID:g60342
R:Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.
J. Virol. 66, 5047-5058, 1992
A:Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 20
C:Superfamily: varicella-zoster virus gene 35 protein

Query Match 78.6%; Score 33; DB 1; Length 303;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPITSC 7
111111
DB 75 CVPITSC 81

Search completed: June 17, 2002, 15:58:25
Job time: 194 sec

ID	PR1	BOVIN	STANDARD;	PRT;	727 AA.
AC	Q28161				
DT	01-MAR-2002	(Rel. 41, Created)			
DT	01-MAR-2002	(Rel. 41, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Plakophilin 1 (Band-6-protein) (B6p).				
GN	PKP1.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_taxid=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=95196971; PubMed=7890138;				
RA	Held H.W., Schmidt A., Zimbelmann R., Schaefer S.,				
RA	Winter-Simanowski S., Stump S., Kelch M., Figeo U., Schnolzer M.,				
RA	Frank W.W.;				
RT	"Cell type-specific desmosomal plaque proteins of the plakoglobin				
RT	family: plakophilin 1 (band 6 protein).";				
RL	Differentiation 58:113-131(1994).				
CC	-1- FUNCTION: SEEMS TO PLAY A ROLE IN JUNCTIONAL PLAQUES (BY				
CC	SIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: Nuclear and associated with desmosomes (By				
CC	similarity).				
CC	-1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.				
CC	-1- SIMILARITY: CONTAINS 8 ARM REPEATS.				
CC	-----				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/isb-sib.ch).				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: Z37975; CAA86029.1; -;				
DR	InterPro: IPR000225; Armadillo.				
DR	Pfam: PF00514; Armadillo_seg; 5.				
DR	SMART: SM00185; ARM; 4.				
DR	PROSITE: PS50176; ARM_REPEAT; 3.				
KW	Cell adhesion; Cytoskeleton; structural protein; Nuclear protein;				
KW	Repeat.				
FT	REPEAT	235	275	ARM 1.	
FT	REPEAT	278	317	ARM 2.	
FT	REPEAT	320	360	ARM 3.	
FT	REPEAT	419	464	ARM 4.	
FT	REPEAT	517	557	ARM 5.	
FT	REPEAT	565	604	ARM 6.	
FT	REPEAT	606	650	ARM 7.	
FT	REPEAT	653	693	ARM 8.	
FT	REPEAT				
SQ	SEQUENCE	727 AA;	80180 MW;	3a27979279BCCEDF	CNC64;
QY	1 CVPPTSC 7				
Db	215 CVPPTSC 221				
Query Match		83.3%;	Score 35;	DB 1;	Length 727;
Best Local Similarity		85.7%;	Pred. No. 15;		
Matches	6; Conservative	0;	Mismatches	1;	Indels
				0;	Gaps
				0;	
RESULT	3				
ID	ZAN_PIG	STANDARD;	PRT;	2476 AA.	
AC	Q28983;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Zocadhesin precursor.				
GN	ZAN.				
OS	Sus scrofa (pig).				

CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX	NCBI_TaxID=9623;
RN	[1]
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 823-830; 859-872; 883-890;
RP	920-925; 960-967; 1235-1244; 1349-1354; 1518-1532; 1624-1656;
RP	1658-1667; 1777-1795 AND 1914-1921.
RC	STRAIN-MEISHAN; TISSUE-Testis;
RX	MEDLINE=96064658; PubMed=7592795;
RA	Hardy D.M., Gaiders D.U.;
RT	"A sperm membrane protein that binds in a species-specific manner to
RT	the egg extracellular matrix is homologous to von Willebrand
RT	factor.";
RL	J. Biol. Chem. 270:26025-26028(1995).
CC	-1- FUNCTION. BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC	OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC	SIGNALING.
CC	-1- SUBUNIT. PROBABLY FORMS COVALENT OLIGOMERS.
CC	-1- SUBCELLULAR LOCATION. TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC	APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
CC	-1- TISSUE SPECIFICITY. IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS,
CC	NOT IN LUNG, LIVER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMUS.
CC	-1- DOMAIN. THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE
CC	ZONA PELLUCIDA.
CC	-1- DOMAIN. DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
CC	THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
CC	SERMATOOCOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
CC	-1- DOMAIN. THE VWFD DOMAINS 2 AND 3 MAY MEDiate COVALENT
CC	OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
CC	-1- PTM: THE MAM DOMAINS AND THE MUCIN-LIKE DOMAINS ARE MISSING
CC	FROM THE ZONADHESIN THAT BINDS TO THE EGG EXTRACELLULAR MATRIX.
CC	PROCESSING MIGHT OCCUR DURING SPERM MATURATION AND/OR
CC	CAPACTATION.
CC	-1- SIMILARITY: CONTAINS 2 MAM DOMAINS.
CC	-1- SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@lsb-sib.ch).

DR	EMBL; U40024; AAC48486.1; .
DR	InterPro: IPRO000561; EGF-like.
DR	InterPro: IPRO000988; MAM.
DR	InterPro: IPRO02919; TIL.
DR	InterPro: IPRO03328; Tila.
DR	InterPro: IPRO01007; VMFC.
DR	InterPro: IPRO01846; Vwd.
DR	Pfam; PF00629; MAM; 2.
DR	Pfam; PF01826; TIL; 5.
DR	Pfam; PF02345; Tila; 5.
DR	Pfam; PF00094; vwd; 4.
DR	SMART; SM00181; EGF; 1.
DR	SMART; SM00137; MAM; 1.
DR	SMART; SK00214; VMC; 2.
DR	SMART; SM00216; VWD; 4.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 4.
DR	PROSITE; PS00740; MAM_1; 1.
DR	PROSITE; PSS0060; MAM_2; 2.
KW	Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW	Repeat.
FT	SIGNAL 1 29 POTENTIAL.
FT	CHAIN 30 2476 ZONADHESIN.
FT	DOMAIN 30 2418 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 2419 2439 POTENTIAL.
FT	DOMAIN 2440 2476 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 31 144 MAM 1.
FT	DOMAIN 147 312 MAM 2.

```

FT DOMAIN 319 687 53 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
FT DOMAIN 688 799 (MOCIN-LIKE DOMAIN).
FT DOMAIN 800 1184 WVED 1 (PARTIAL).
FT DOMAIN 1185 1573 WVED 2.
FT DOMAIN 1574 1968 WVED 3.
FT DOMAIN 1969 2370 WVED 4.
FT DOMAIN 2366 2402 WVED 5.
FT DISULFID 2370 2381 EGF-LIKE.
FT DISULFID 2375 2390 BY SIMILARITY.
FT DISULFID 2392 2401 BY SIMILARITY.
FT CARBOHYD 109 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 758 758 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 833 833 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1154 1154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1329 1329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1448 1448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1544 1544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1596 1596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1654 1654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1843 1843 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1965 1965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2122 2122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2165 2165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2178 2178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2329 2329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2359 2359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 823 823 C -> V (IN REF. 1; AA SEQUENCE).
FT CONFLICT 923 923 S -> Y (IN REF. 1; AA SEQUENCE).
FT CONFLICT 965 965 W -> Y (IN REF. 1; AA SEQUENCE).
FT CONFLICT 1241 1241 S -> K (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 2476 AA: 270364 MW: A13B890375A6348C CRC64;

```

```

Query Match 81.0%; Score 34; DB 1; Length 2476;
Best Local Similarity 71.4%; Pred. No. 73;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 CVPITSC 7
Db 1505 CVPITSC 1511

RESULT 4
ID UL24_HSVSA STANDARD; PRT; 303 AA.
AC G01005;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Gene 20 protein.
GN 20.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wiltmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Hoess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL24,
CC EHV-1 37, EBV BXRFL, HCMV UL76, ILTV ORF3, AND VZV 35.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64346; CAA45644.1; -
DR PIR: B36808; W2BEM6.
DR InterPro: IPR002580; Herpes_UL24.
DR Pfam: PF01646; Herpes_UL24; 1.
SQ SEQUENCE 303 AA: 34942 MW: DF6D59F7A1C83A0B CRC64;

```

```

Query Match 78.6%; Score 33; DB 1; Length 303;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 CVPITSC 7
Db 75 CVPITSC 81

```

```

RESULT 5
ID FBX3_HUMAN STANDARD; PRT; 471 AA.
AC Q9UR99; Q9NUX2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE F-box only protein 3.
GN FBX3 OR FBX3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunaga H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 6-415 FROM N.A.
RX MEDLINE=20003061; PubMed=10531037;
RA Winston J.T., Koepf D.M., Zhu C., Elledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins.";
RL Curr. Biol. 9:1180-1182(1999).
CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK001943; BA91991.1; -
DR EMBL: AF176702; AAF03702.1; -
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; F-box; 1.
DR PROSITE: PS50181; F-box; 1.
KW Ubiquitin conjugation.
FT DOMAIN 10 56 F-BOX.
FT DOMAIN 419 452 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 453 456 POLY-ARG.
FT DOMAIN 463 466 POLY-ARG.

```

FT CONFLICT 164 164 T -> A (IN REF. 2).
 FT CONFLICT 414 415 EM -> VS (IN REF. 2).
 SQ SEQUENCE 471 AA; 54590 MW; F7AA88193E14E67E CRC64;

Query Match 78.6%; Score 33; DB 1; Length 471;
 Best Local Similarity 71.4%; Pred. No. 24;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVPJLSC 7
 Db 193 CLPLTFC 199

RESULT 6

KE4L_CAEEL STANDARD; PRT; 515 AA.
 AC 09XT07;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN H13N06.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chordata; Rhabditiida; Rhabditoidea;
 OC Rhabditiida; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=62339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Leonard N.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KE4/CAT5P FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 299942; CAB17070.1; -
 DR WormPep; H13N06.5; CE18815.
 DR InterPro: IPR002395; Kintinogen.
 DR InterPro: IPR003689; Zip.
 DR Pfam; PF02535; Zip; 1.
 DR PRINTS; PR00334; KININOGEN.
 KW Hypothetical protein; Transmembrane; Glycoprotein.
 FT TRANSMEM 27 47
 FT TRANSMEM 49 69
 FT TRANSMEM 214 234
 FT TRANSMEM 247 267
 FT TRANSMEM 297 317
 FT TRANSMEM 386 406
 FT TRANSMEM 429 449
 FT TRANSMEM 463 483
 FT DOMAIN 92 182
 FT CARBOHYD 7 7
 FT CARBOHYD 237 237
 FT CARBOHYD 379 379
 FT CARBOHYD 488 488
 SQ SEQUENCE 515 AA; 55500 MW; 17D7E854FAE1DAF CRC64;

Query Match 78.6%; Score 33; DB 1; Length 515;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CVPJLSC 7
 Db 24 CLPLTSC 30

RESULT 7

ZAN_RABIT STANDARD; PRT; 2282 AA.
 ID ZAN_RABIT
 AC P57999;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Zonadhesin (Fragment).
 GN ZAN.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Lea I.A., Sivasubramanian P., Richardson R.T., O'Rand M.G.;
 RT "Sequence of rabbit zonadhesin."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING (BY SIMILARITY).
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
 CC -1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE
 CC ZONA PELLUCIDA.
 CC -1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
 CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
 CC SPERMATOZOEA OR PROMOTING ADHESION TO THE OVUDUCTAL ISTHMUS.
 CC -1- DOMAIN: THE VWFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT
 CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
 CC -1- SIMILARITY: CONTRAINS AT LEAST 2 MAM DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF244982; AAF63342.2; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR000996; MAM.
 DR InterPro: IPR002919; TIL.
 DR InterPro: IPR003328; TILA.
 DR InterPro: IPR001007; VWFC.
 DR InterPro: IPR001846; Vwd.
 DR Pfam; PF00629; MAM; 2.
 DR Pfam; PF01826; TIL; 5.
 DR Pfam; PF02345; TILA; 5.
 DR Pfam; PF02094; Vwd; 4.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00137; MAM; 1.
 DR SMART; SM00214; VNC; 5.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS00740; MAM_1; FALSE_NEG.
 DR PROSITE; PS50060; MAM_2; 2.
 KW Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 KW Repeat.
 FT NON_TER 1 1
 FT DOMAIN <1 2235
 FT TRANSMEM 2236 2256
 FT DOMAIN 2257 2282
 FT DOMAIN <1 147
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 MAM 1.

FT DOMAIN 150 315 MAM 2.
 FT DOMAIN 315 498 26 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
 FT DOMAIN 499 610 (MUCIN-LIKE DOMAIN).
 FT DOMAIN 611 995 WVED 1. (PARTIAL).
 FT DOMAIN 996 1384 WVED 2.
 FT DOMAIN 1385 1787 WVED 3.
 FT DOMAIN 1788 2189 WVED 4.
 FT DOMAIN 2189 2221 WVED 5.
 FT DISULFID 2189 2200 EGF-LIKE.
 FT DISULFID 2194 2209 BY SIMILARITY.
 FT DISULFID 2211 2220 BY SIMILARITY.
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1141 1141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1259 1259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1270 1270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1355 1355 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1467 1467 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1483 1483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1662 1662 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1997 1997 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2178 2178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 2282 AA; 248290 MW; 380FA81093454892 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 2282;
 Best Local Similarity 71.4%; Pred. No. 1; e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVPETSC 7
 DB 1316 CVPINOC 1322

RESULT 8
 HMCT_BOMMO STANDARD; PRT: 3133 AA.
 ID HMCT_BOMMO
 AC P80992;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hemocytin precursor (Humoral lectin).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_Taxid=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FUJOU X TOKAI; TISSUE=Hemocyte;
 RX MEDLINE=95178544; PubMed=7873598;
 RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Tanaka K., Kadono-Okuda K., Kato Y., Mori H.;
 RT "Cloning and expression of the gene of hemocytin, an insect humoral
 lectin which is homologous with the mammalian von Willebrand
 factor.";
 RT factor.";
 RL Blochim. Biophys. Acta 1260:245-258(1995).
 RN [2]
 RP SEQUENCE OF 2221-3133 FROM N.A.
 RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Yamakawa M.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases
 CC -1- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR
 CC ENCAPSULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL
 CC METAMORPHOSIS.
 CC -1- INDUCTION: HEMAGGLUTINATION ACTIVITY IS INCREASED BY BACTERIAL
 CC OR VIRAL INFECTION AND INHIBITED BY D-MANNOSE, N-ACETYL-D-
 CC GALACTOSAMINE AND D-MALTOSE.
 CC -1- PTM: MAY BE CONVERTED INTO THE 260 kDa MATURE HEMOCYTIN BY

CC PROTEOLYSIS.
 CC -1- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.
 CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
 CC OF HUMAN MUCIN 2.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 WFEC DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.stb.ch).
 CC -----
 CC EMBL: D29738; BAA06160.1; -
 CC EMBL: D14035; BAA03124.1; -
 CC InterPro: IPR000359; Cys_knot.
 CC InterPro: IPR000421; FA58_C.
 CC InterPro: IPR002172; LDL_recept_A.
 CC InterPro: IPR002919; TIL.
 CC InterPro: IPR001007; WFEC.
 CC InterPro: IPR001846; Wvd.
 CC InterPro: IPR001878; Znf_CCHC.
 CC InterPro: IPR000083; fibronectin_type_1.
 CC Pfam: PF00754; F5_F8_type_C; 2.
 CC Pfam: PF01826; TIL; 6.
 CC Pfam: PF00094; Wvd; 3.
 CC SMART: SM00041; CT; 1.
 CC SMART: SM00231; FA58C; 2.
 CC SMART: SM00058; FN1; 1.
 CC SMART: SM00192; LDla; 1.
 CC SMART: SM00214; WVC; 3.
 CC SMART: SM00216; WVD; 3.
 CC SMART: SM00343; Znf_C2HC; 1.
 CC PROSITE: PS01185; CTCK_1; 1.
 CC PROSITE: PS01208; WFEC; FALSE_NEG.
 CC PROSITE: PS01225; CTCK_2; 1.
 CC PROSITE: PS01285; FA58C_1; 2.
 CC PROSITE: PS01286; FA58C_2; 2.
 CC Lectin; Glycoprotein; Signal; Repeat; Cell adhesion.
 CC KW SIGNAL
 CC FT CHAIN ? 3133 ?
 CC FT DOMAIN 29 131
 CC FT DOMAIN 153 240
 CC FT DOMAIN 248 613
 CC FT DOMAIN 940 1095
 CC FT DOMAIN 1116 1254
 CC FT DOMAIN 1283 1356
 CC FT DOMAIN 1620 1951
 CC FT DOMAIN 1952 2315
 CC FT DOMAIN 2230 2321
 CC FT DOMAIN 2335 2361
 CC FT DOMAIN 2435 2469
 CC FT DOMAIN 2553 2622
 CC FT DOMAIN 2842 2907
 CC FT DOMAIN 2971 3076
 CC FT DOMAIN 895 914
 CC FT DOMAIN 1267 1270
 CC FT DOMAIN 1425 1428
 CC FT DOMAIN 1447 1450
 CC FT DOMAIN 1474 1479
 CC FT DOMAIN 2148 2153
 CC FT DOMAIN 2156 2159
 CC FT DOMAIN 2341 2344
 CC FT DISULFID 940 1095
 CC FT DISULFID 1116 1254
 CC FT DISULFID 2991 3040
 CC FT DISULFID 2991 3054
 CC FT DISULFID 3004 3070
 CC FT DISULFID 3020 3072
 CC FT DISULFID ? 3075
 CC
 CC B1.
 CC B2.
 CC WFEC 1.
 CC WFEC 2.
 CC CTCK.
 CC POLY-THR.
 CC POLY-GLU.
 CC POLY-THR.
 CC POLY-SER.
 CC POLY-PRO.
 CC POLY-PRO.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.

```

FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1170 1170 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1622 1622 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1727 1727 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1847 1847 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1985 1985 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2093 2093 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2113 2113 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2161 2161 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2276 2276 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2451 2451 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2647 2647 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2654 2654 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2663 2663 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2794 2794 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2810 2810 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2865 2865 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2929 2929 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 2964 2964 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 3028 3028 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT VARIANT 1288 1288 R -> G.
FT VARIANT 1305 1305 T -> S.
SQ SEQUENCE 3133 AA: 343350 MW: E5210D5D14A7B2B2 CRC64:

```

```

Query Match 78.6% Score 33; DB 1; Length 3133;
Best Local Similarity 71.4%; Pred. NO. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CVP1ISC 7
    11111
Db 203 CVPATQC 209

```

```

RESULT 9
MUC2_HUMAN STANDARD; PRT; 5179 AA..
ID MUC2_HUMAN
AC Q02817; Q14878;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mucin 2 precursor (Intestinal mucin 2).
GN MUC2 OR SMOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=94132002; PubMed=8300571;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
RT "Molecular cloning of human intestinal mucin (MUC2) cDNA.
RT Identification of the amino terminus and overall sequence similarity
RT to prepro-von Willebrand factor."
RL J. Biol. Chem. 269:2440-2446(1994).
RN [2]
RP SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=93016075; PubMed=1400449;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Rothe E.-M., Lagace R.E.,
RA Kim Y.S.;
RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT both upstream and downstream of its central repetitive region."
RL J. Biol. Chem. 267:21375-21383(1992).
RN [3]
RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RX MEDLINE=91358717; PubMed=1885763;
RA Toribara N.W., Gum J.R. Jr., Culhane P.J., Lagace R.E., Hicks J.W.,

```

```

RA Petersen G.M., Kim Y.S.;
RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT and polymorphism."
RL J. Clin. Invest. 88:1005-1013(1991).
CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
CC OTHER MUCOSAL MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
CC AGENTS AT MUCOSAL SURFACES.
CC -1- SUBUNIT: MULTIMERIC.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC BRONCHUS, CERVIX AND GALL BLADDER.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC VARIES AMONG DIFFERENT ALLELES.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
CC OF SILKWORM HEMOCYTIN.
CC -1- SIMILARITY: CONTAINS 2 WMFC DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L21998; AAB95295.1; -
DR EMBL: M74027; AAA59875.1; -
DR EMBL: M94131; AAA59163.1; -
DR EMBL: M94132; AAA59164.1; -
DR MIM: 158370; -
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002400; GF_Cys_knot.
DR InterPro: IPR001007; WMFC.
DR InterPro: IPR001846; Vwd.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00094; Vwd; 4.
DR PRINTS: PR00438; GFCSKNOT.
DR SMART: SM00214; VWC; 2.
DR SMART: SM00216; VWD; 4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS01208; WMFC; 2.
KW Glycoprotein; Repeat; Signal.
FT SIGNAL 1 20
FT CHAIN 21 5179
FT DOMAIN 1401 1747
FT REPEAT 1401 1416
FT REPEAT 1417 1432
FT REPEAT 1433 1448
FT REPEAT 1449 1464
FT REPEAT 1465 1471
FT REPEAT 1472 1478
FT REPEAT 1479 1494
FT REPEAT 1495 1517
FT REPEAT 1518 1533
FT REPEAT 1534 1556
FT REPEAT 1557 1572
FT REPEAT 1573 1596
FT REPEAT 1597 1612
FT REPEAT 1613 1635
FT REPEAT 1636 1651
FT REPEAT 1652 1675
FT REPEAT 1676 1683
FT REPEAT 1684 1699
FT REPEAT 1700 1715
FT REPEAT 1716 1731
FT REPEAT 1732 1747
FT REPEAT 1747 1747

```

```

FT DOMAIN 4815 4886 WPMC 1.
FT DOMAIN 4924 4991 WPMC 2.
FT DOMAIN 5075 5160 CHECK.
FT DISULFID 5075 5122 BY SIMILARITY.
FT DISULFID 5089 5136 BY SIMILARITY.
FT DISULFID 5098 5152 BY SIMILARITY.
FT DISULFID 5102 5154 BY SIMILARITY.
FT DISULFID 5102 5154 BY SIMILARITY.
FT DISULFID 5102 5154 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1134 1134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1215 1215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1230 1230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1787 1787 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1820 1820 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4339 4339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4351 4351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4362 4362 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4373 4373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4422 4422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4438 4438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4502 4502 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 1504 1504 G -> S (IN REF. 2).
FT CONFLICT 4192 4192
SQ SEQUENCE 5179 AA; 540295 MW; 85CD7571FB9A5663 CRC64;

```

```

Query Match Score 33; DB 1; Length 5179;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 CVPITSC 7
DB 712 CVPITSC 718

```

```

RESULT 10
PRTF_BOMMO STANDARD: PRT: 77 AA.
AC Q10731;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Fungal protease inhibitor F precursor (FPI-F).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96424990; PubMed=8827441;
RA Pham T.N., Hayashi K., Takano R., Nakazawa H., Mori H., Ichida M.,
RA Itoh M., Eguchi M., Matsubara F., Hara S.;

```

```

RT "Expression of Bombyx family fungal protease inhibitor F from Bombyx
RT mori by baculovirus vector."
RT J. Blochem. 119:1080-1085(1996).
RN [2]
RN SEQUENCE FROM N.A.
RN STRAIN=C124; TISSUE=Fat body;
RA Itoh M., Takenaka T., Ashikari T., Eguchi M.;
RT "CDNA cloning and expression of a novel type protease inhibitor
RT (PPI-F) from the silkworm, Bombyx mori."
RT J. Sericultural Sci. Jpn. 65:326-333(1996).
RN [3]
RN SEQUENCE OF 23-77.
RX MEDLINE=96271002; PubMed=8830035;
RA Pham T.-N., Hayashi K., Takano R., Itoh M., Eguchi M., Shibata H.,
RA Tanaka T., Hara S.;
RT "A new family of serine protease inhibitors (Bombyx family) as
RT established from the unique topological relation between the
RT positions of disulphide bridges and reactive site."
RT J. Biochem. 119:428-434(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
CC EMBL: S83181; AAR46908.1; -
CC EMBL: D38075; BAA22409.1; -
CC InterPro: IPR002919; TIL.
DR Pfam: PF01826; TIL; 1.
KW Serine protease inhibitor; Signal.
FT CHAIN 1 22
FT ACT_SITE 23 77
FT DISULFID 51 52
FT DISULFID 25 57
FT DISULFID 36 49
FT DISULFID 40 77
FT DISULFID 59 71
SQ SEQUENCE 77 AA; 8492 MW; B9CFC085DDA10354 CRC64;

```

```

Query Match Score 32; DB 1; Length 77;
Best Local Similarity 57.1%; Pred. No. 6.9;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 CVPITSC 7
DB 71 CVPITSC 77

```

```

RESULT 11
PRTF_PAT STANDARD: PRT: 305 AA.
AC P24392; Q63733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE peroxisome assembly factor-1 (PAF-1) (Peroxin-2) (Peroxisomal membrane
DE protein 3).
GN PMP35 OR PAF1 OR PMP35.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=91163637; PubMed=1750930;
RA Tsukamoto T., Miura S., Fujiki Y.;
RT "Restoration by a 35k membrane protein of peroxisome assembly in a
RT peroxisome-deficient mammalian cell mutant."

```

```

RL Nature 350:77-81(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY: TISSUE-Liver;
RX MEDLINE=94309666; PubMed=8035823;
RA Tsukamoto T., Shimozawa N., Fujiki Y.;
RT Peroxisome assembly factor 1: nonsense mutation in a peroxisome-
RL deficient Chinese hamster ovary cell mutant and deletion analysis.";
MOI. Cell. Biol. 14:5458-5465(1994).
CC -1- FUNCTION: SOMEWHAT IMPLICATED IN THE BIOGENESIS OF PEROXISOMES.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57988; CAA41054.1; -
DR EMBL; D30616; BAA06306.1; -
DR EMBL; D30617; BAA06307.1; -
DR PIR; S14290; S14290.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Transmembrane; Peroxisome; Zinc-finger.
FT TRANSMEM 140 159
FT TRANSMEM 195 213 POTENTIAL.
FT ZN_RING 284 284 RING-TYPE.
FT VARIANT 125 125 R -> P.
SQ SEQUENCE 305 AA; 34767 MW; 249058A5A5A8E854 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 305;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPRTS 6
DB 224 CIPRTS 229

RESULT 12
PSPB_DICDI STANDARD; PRT; 379 AA.
AC P54704;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Prespore protein B precursor.
GN PSPB.
OS Dictyostelium discoideum (Slime mold)
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AX3;
RX MEDLINE=94326657; PubMed=8050366;
RA Powell-Coffman J.A., Firtel R.A.;
RT "Characterization of a novel Dictyostelium discoideum
RT prespore-specific gene, PSPB, reveals conserved regulatory
RT sequences.";
RL Development 120:1601-1611(1994).
CC -1- SIMILARITY: CONTAINS 3 PRESPORE MOTIFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S72639; AAC60506.1; -
DR Dictydb; DD02053; PSPB.
DR InterPro; IPR003645; FOLN.
DR SMART; SM00274; FOLN; 1.
DR GLYCOPROTEIN; Repeat; Sporulation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 379
FT REPEAT 224 236 PRESPORE PROTEIN B.
FT REPEAT 289 301 PRESPORE MOTIF 1.
FT REPEAT 319 331 PRESPORE MOTIF 2.
FT REPEAT 319 331 PRESPORE MOTIF 3.
SQ SEQUENCE 379 AA; 42719 MW; EA92A6ED69D05431 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 379;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPRTSC 7
DB 283 CIPRTTC 289

RESULT 13
PMT7_BOVIN STANDARD; PRT; 491 AA.
AC Q06154;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanocyte protein Pmel 17 (Retinal pigment epithelial-specific
DE protein) (Fragment).
GN SILV OR PMEL17 OR RPE1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Retina;
RX MEDLINE=93122163; PubMed=1478275;
RA Kim R.Y., Wistow G.J.;
RT "The CDNA RPE1 and monoclonal antibody HMB-50 define gene products
RT preferentially expressed in retinal pigment epithelium.";
RL Exp. Eye Res. 55:657-662(1992).
CC -1- FUNCTION: COULD BE A MELANOCYTOGENIC ENZYME (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: RETINAL PIGMENT EPITHELIUM.
CC -1- SIMILARITY: BELONGS TO THE PMEL-17/MMB FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81193; AAA30419.1; -
DR InterPro; IPR00601; PKD_domain.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
DR Transmembrane; Glycoprotein; Melanin biosynthesis; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 423 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 424 444

```

FT DOMAIN 445 491 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 60 150 PKD.
FT DOMAIN 148 256 8 X 13 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 148 160 1.
FT REPEAT 161 173 2.
FT REPEAT 174 186 3.
FT REPEAT 187 199 4.
FT REPEAT 200 212 5.
FT REPEAT 213 225 6.
FT REPEAT 232 243 7.
FT REPEAT 244 256 8.
FT DOMAIN 304 394 CYS-RICH.
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 491 AA; 51669 MW; 28BF55DFBD397D6D CRC64;

Query Match Best Local Similarity 76.28; Score 32; DB 1; Length 491;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPLTSC 7
Db 139 IPLTSC 144

RESULT 14
PM17_HUMAN STANDARD: PRF: 661 AA.
AC P40967; Q16565; Q14817; Q12763; Q14448;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (Melanocyte lineage-specific antigen gp100) (Melanoma-associated ME20 antigen) (ME20M/ME20S) (ME20-M/ME20-S) (95 kDa melanocyte-specific secreted glycoprotein).
GN SILV OR PHE17 OR D1253E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92021023; Pubmed=1924386;
RA Kwon B.S., Chintamaneni C., Kozak C.A., Copeland N.G., Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y., Kim K.-K.;
RT "A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on mouse chromosome 10 and is in a syntenic region on human chromosome 12.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94327568; Pubmed=7519602;
RA Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Fidor C.G.;
RT "Molecular characterization of the melanocyte lineage-specific antigen gp100.";
RL J. Biol. Chem. 269:20126-20133(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154052; Pubmed=8592076;
RA Baillin T., Lee S.T., Spritz R.A.;
RT "Genomic organization and sequence of D1253E (Pmel 17), the human homologue of the mouse silver (si) locus.";
RL J. Invest. Dermatol. 106:24-27(1996).
RN [4]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 25-53.
RX MEDLINE=94235165; Pubmed=8179825;
RA Maresh G.A., Marken J.S., Neubauber M., Aruffo A., Hellstrom I., Hellstrom K.E., Marguardt H.;
RT "Cloning and expression of the gene for the melanoma-associated ME20 antigen.";
RL DNA Cell Biol. 13:87-95(1994).

RN [5]
RP SEQUENCE FROM N.A.
RA Kwon B.S., Kim K.-K., Heng H.H., Shi X.M., Teul L., Lee Z.H., Yoon B., Pickard R.T.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Vogel A.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: COULD BE A MELANOGENIC ENZYME. COULD REPRESENT AN ONCOGENIC SELF-ANTIGEN THAT IS NORMALLY EXPRESSED AT LOW LEVELS IN QUIESCENT ADULT MELANOCYTES BUT OVEREXPRESSED BY PROLIFERATING NEONATAL MELANOCYTES AND DURING TUMOR GROWTH. RELEASE OF THE SOLUBLE FORM, ME20-S, COULD PROTECT TUMOR CELLS FROM ANTIBODY MEDIATED IMMUNITY.
CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL). THERE IS ALSO A SECRETED SOLUBLE FORM, ME20-S, PROBABLY PRODUCT OF PROTEOLYTIC CLEAVAGE.
CC -! TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOMAS. SOME EXPRESSION WAS FOUND IN DYSPLASTIC NEVI. NOT FOUND IN NORMAL TISSUES NOR IN CARCINOMAS.
CC -! SIMILARITY: BELONGS TO THE PHE17/17MMB FAMILY.
CC -! SIMILARITY: CONTAINS 1 PKD DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL: M7348; AAA60121.1; -
DR EMBL: S7303; AAC60634.1; -
DR EMBL: U31799; AAB00386.1; -
DR EMBL: U31807; AAB00386.1; JOINED.
DR EMBL: U31807; AAB00386.1; JOINED.
DR EMBL: U31797; AAB00386.1; JOINED.
DR EMBL: U31798; AAB00386.1; JOINED.
DR EMBL: U01874; AAA18479.1; -
DR EMBL: U20093; AAB19181.1; -
DR EMBL: U19491; AAB19181.1; JOINED.
DR EMBL: M32295; AAA35930.1; ALT_INIT.
DR MIM: 155550; -
DR InterPro: IPR000601; PKD_domain.
DR Pfam: PF00801; PKD_1.
DR SMART: SM00089; PKD_1.
DR PROSITE: PS50093; PKD_1.
KW Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat; Antigen.
FT SIGNAL 1 24
FT CHAIN 25 661
FT DOMAIN 25 595
FT TRANSMM 596 616
FT DOMAIN 617 661
FT DOMAIN 255 292
FT DOMAIN 315 444
FT REPEAT 315 327
FT REPEAT 328 340
FT REPEAT 341 353
FT REPEAT 354 366
FT REPEAT 367 379
FT REPEAT 380 392
FT REPEAT 393 405
FT REPEAT 406 418
FT REPEAT 419 431
FT REPEAT 432 444
FT CARBOHD 81 81
FT CARBOHD 106 106
FT CARBOHD 111 111
FT CARBOHD 321 321
FT CARBOHD 568 568
FT CONFLICT 274 274
CC MELANOCYTE PROTEIN PHE17.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC 10 X 13 AA APPROXIMATE TANDEM REPEATS.
CC 1.
CC 2.
CC 3.
CC 4.
CC 5.
CC 6.
CC 7.
CC 8.
CC 9.
CC 10.
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC L -> P (IN REF. 1 AND 5).

FT CONFLICT 587 587 P -> PVPGLILT (IN REF. 1).
 FT CONFLICT 592 592 G -> GG (IN REF. 4).
 FT CONFLICT 597 597 P -> R (IN REF. 1).
 FT CONFLICT 642 661 RIFCSPIGENSPILSGOV -> ASSALVPLVRIAPSSVG
 FT SEQUENCE 661 AA; 70255 MW; 8A904FAB16715653 CRC64;
 SQ
 Query Match 76.2%; Score 32; DB 1; Length 661;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VPLTSC 7
 Db 296 IRLTSC 301
 RESULT 15
 SIF2_DROME STANDARD; PRT; 2044 AA.
 ID SIF2_DROME
 AC P1620;
 DT 15-JUN-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Still life protein type 2 (SIF type 2).
 GN SIF.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Head;
 RX MEDLINE=97153054; PubMed=8999801;
 RA Sone M., Hoshino M., Suzuki E., Kuroda S., Kabuchi K., Nakagoshi H.,
 RA Saigo K., Nabeshima Y.-I., Hama C.;
 RT Still life, a protein in synaptic terminals of Drosophila homologous
 RT to GDP-GTP exchangers.;
 RL Science 275:543-547(1997).
 RN [2]
 RP ERRATUM.
 RA Sone M., Hoshino M., Suzuki E., Kuroda S., Kabuchi K., Nakagoshi H.,
 RA Saigo K., Nabeshima Y.-I., Hama C.;
 RL Science 275:1405-1405(1997).
 CC -1- FUNCTION: REGULATES SYNAPTIC DIFFERENTIATION THROUGH THE
 CC ORGANIZATION OF ACTIN CYTOSKELETON POSSIBLY BY ACTIVATING RHO-LIKE
 CC GTPASES. IS LIKELY A FACTOR IN THE CASCADE OF RAC1 OR CDC42 IN THE
 CC NEURONS.
 CC -1- SUBCELLULAR LOCATION: LOCALIZES TO THE SUBMEMBRANOUS REGION OF
 CC SYNAPTIC TERMINALS.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SIF TYPE 1 (AC P1621) AND SIF
 CC TYPE 2 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DEVELOPMENTAL STAGE: AT STAGE 14, EXPRESSION OCCURS IN EACH
 CC SEGMENT OF THE CENTRAL NERVOUS SYSTEM. AT STAGE 17, EXPRESSION
 CC BECOMES RESTRICTED TO THE SYNAPTIC REGIONS OF THE BRAIN AND
 CC VENTRAL NERVE CORD, WHERE SYNAPSES UNDERGO MATURATION.
 CC -1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
 CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D86546; BAA13108.1; -;
 CC DR HSSP: P08567; IPI.
 CC DR FlyBase: FBgn0019652; sif.
 CC DR InterPro: IPR001331; GDS_CDC24.

DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR003116; RBD.
 DR InterPro: IPR000219; RhogEF.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00169; PH; 2.
 DR Pfam: PF02196; RBD; 1.
 DR Pfam: PF00621; RhogEF; 1.
 DR SMART: SM00238; PDZ; 1.
 DR SMART: SM00233; PH; 2.
 DR SMART: SM00455; RBD; 1.
 DR SMART: SM00325; RhogEF; 1.
 DR PROSITE: PS50010; DH_2; 1.
 DR PROSITE: PS00741; DH_1; 1.
 DR PROSITE: PS50106; PDZ; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW Guanine-nucleotide releasing factor; Developmental protein; Synapse;
 KW Repeat; Alternative splicing.
 FT DOMAIN 62 249 4 X 25 AA APPROXIMATE REPEAT.
 FT REPEAT 62 86 1.
 FT REPEAT 94 118 2.
 FT REPEAT 154 178 3.
 FT REPEAT 225 249 4.
 FT DOMAIN 819 937 PH 1.
 FT DOMAIN 1184 1273 PDZ.
 FT DOMAIN 1408 1602 DH.
 FT DOMAIN 1674 1767 PH 2.
 FT DOMAIN 467 470 POLY-PRO.
 FT DOMAIN 646 649 POLY-ARG.
 FT DOMAIN 1295 1298 POLY-PRO.
 FT DOMAIN 1898 1909 POLY-GLN.
 FT DOMAIN 1929 1933 POLY-PRO.
 SQ SEQUENCE 2044 AA; 228324 MW; 75D7CF21F49654B6 CRC64;
 QY 1 CVPLTSC 7
 Db 502 CAPLISC 508
 Query Match 76.2%; Score 32; DB 1; Length 2044;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Search completed: June 17, 2002, 16:24:30
 Job time: 1559 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:40 ; Search time 172.85 seconds
(without alignments)
7.006 Million cell updates/sec

Title: US-09-761-636a-11

Perfect score: 42
Sequence: 1 CVPLTSC 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	39	92.9	2543	5	P90649
2	38	90.5	127	6	O9GMY9
3	36	85.7	337	5	O18464
4	36	85.7	980	2	O24780
5	36	85.7	2533	5	O27183
6	35	85.7	2533	5	P90589
7	35	83.3	193	13	O90993
8	35	83.3	389	6	O97887
9	35	83.3	462	11	O9CXM8
10	35	83.3	727	6	O28161
11	35	83.3	1637	6	O9XSV8
12	35	83.3	3680	5	O9VR08
13	34	81.0	88	5	O9VJX0
14	34	81.0	1506	13	P79927
15	34	81.0	1513	11	O62635
16	34	81.0	2843	4	O9Y6R7

17	33	78.6	55	10	O38989
18	33	78.6	101	12	O91F35
19	33	78.6	143	2	O9ACU2
20	33	78.6	191	6	O28583
21	33	78.6	196	11	O9D226
22	33	78.6	207	16	O25713
23	33	78.6	207	16	O92M78
24	33	78.6	220	5	O9NDK2
25	33	78.6	223	11	O62220
26	33	78.6	230	11	O64507
27	33	78.6	303	12	O40640
28	33	78.6	303	12	O92585
29	33	78.6	367	11	O91V19
30	33	78.6	415	11	O91E4
31	33	78.6	427	4	O96Q28
32	33	78.6	471	4	O9H0V2
33	33	78.6	480	11	O9DC63
34	33	78.6	480	11	O9D2J5
35	33	78.6	644	10	O9FKV9
36	33	78.6	644	10	O944Q4
37	33	78.6	647	5	O21573
38	33	78.6	891	10	O9SHG9
39	33	78.6	924	10	O9FSP9
40	33	78.6	1227	10	O9LXP4
41	32	76.2	66	13	O9PTM2
42	32	76.2	68	6	O97500
43	32	76.2	68	13	O9PTM3
44	32	76.2	75	6	O18843
45	32	76.2	85	4	O9P1K3

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	2543 AA.
P90649	AC	P90649		
DR	01-MAY-1997	(TREMBLrel. 03, Created)		
DT	01-MAY-1997	(TREMBLrel. 03, Last sequence update)		
DE	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	156D SURFACE ANTIGEN.			
OC	Parametecium primaurella.			
OC	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;			
OC	Parametecium.			
OX	NCBI_TaxID=5886;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=156;			
RX	MEDLINE=96313351; PubMed=8768434;			
RA	Bourgain-Guglielmetti F., Caron;			
RT	"Molecular characterization of the D surface protein gene subfamily 1n			
RT	Parametecium primaurella."			
RL	J. Eukaryot. Microbiol. 43:303-314(1996).			
DR	EMBL: X9616; CAA65436.1; -			
DR	InterPro: IPR002223; Kunitz-BPTI.			
DR	InterPro: IPR002895; Parametecium_SA.			
DR	Pfam: PF01508; Parametecium_SA. 19.			
KW	PRINTS: PR00759; BASICTPASE.			
KW	Serine protease inhibitor.			
SO	SEQUENCE 2543 AA; 267041 MW; 828EF797CB012902 CRC64;			

Query Match 92.9%; Score 39; DB 5; Length 2543;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTSC 7
DB 2156 CIPITSC 2164

RESULT 2

Q9GMW9
ID Q9GMW9 PRELIMINARY; PRT; 127 AA.
AC Q9GMW9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 13.0 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN PARIENTAL LOBE;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047606; BAB12130.1; -
KW Hypothetical protein.
SQ SEQUENCE 127 AA; 12961 MW; EEl68525256992C0 CRC64;

Query Match 90.5%; Score 38; DB 6; Length 127;
Best Local Similarity 85.7%; Pred. No. 1.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPJTSC 7
Db 107 CVPJTSC 113

RESULT 3
ID 018464 PRELIMINARY; PRT; 337 AA.
AC 018464;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 35.6 KDA PROTEIN PRECURSOR.
GN HMEGFL-1.
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Herdmania.
OX NCBI_TaxID=7733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CURVATA;
RA Arnold J.M., Eri R., Lavlin M.F.;
RT "A novel gene containing multiple EGF-like motifs transiently
RT expressed in the papillae of the ascidian tadpole larvae";
RL Dev. Dyn. 0:0-0(1997).
DR EMBL; U82540; AA667704.1; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002919; TIL.
DR Pfam; PF01826; TIL; 4.
DR SMART; SM00181; EGF; 4.
KW Signal; Hypothetical protein.
FT SIGNAL
FT CHAIN 1 18 POTENTIAL.
SQ SEQUENCE 337 AA; 35617 MW; SCCA0924118DFC6 CRC64;

Query Match 85.7%; Score 36; DB 5; Length 337;
Best Local Similarity 71.4%; Pred. No. 7.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPJTSC 7
Db 199 CVPJTSC 205

RESULT 4
ID 024780 PRELIMINARY; PRT; 980 AA.
AC 024780;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CELLODEXTRIN PHOSPHORYLASE.
GN CDP.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27405;
RA Kawaguchi T., Ikeuchi Y., Tsutsumi N., Kan A., Sumitani J., Arai M.;
RT "Cloning, nucleotide sequence, and expression of the Clostridium
RT thermocellum cellodextrin phosphorylase gene and its application to
RT synthesis of cellulase inhibitors";
RL J. Ferment. Bioeng. 85:144-149(1998).
DR EMBL; AB006822; BAA22081.1; -
SQ SEQUENCE 980 AA; 111183 MW; 8A8C8EE3F5370831 CRC64;

Query Match 85.7%; Score 36; DB 2; Length 980;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPJTSC 7
Db 550 CVPJTSC 556

RESULT 5
ID 027183 PRELIMINARY; PRT; 2533 AA.
AC 027183;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-5ID-IMMOBILIZATION ANTIGEN.
GN ALPHA-5ID-GENE.
OS Paramesicium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Paramesicium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=51;
RA Schmidt H.J.;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X85135; CA59447.1; -
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002895; Paramesicium_SA.
DR Pfam; PF01508; Paramesicium_SA; 22.
DR PRINTS; PR00759; BASICPTASE.
KW Serine protease inhibitor.
SQ SEQUENCE 2533 AA; 263996 MW; 261BD09806BC344D CRC64;

Query Match 85.7%; Score 36; DB 5; Length 2533;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPJTSC 7
Db 2149 CVPJTSC 2155

RESULT 6
P90589

ID P90589 PRELIMINARY; PRT; 2533 AA.
 AC P90589;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALPHA-51D IMMOBILIZATION ANTIGEN.
 GN ALPHA-51D.
 OS Paramesicium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 OC Paramesicium.
 OX NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=51;
 RA Schegemann K., Klein H., Schmidt H.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X66400; CAA65264.1; -
 DR InterPro: IPR002223; Kunitz_BPTI
 DR InterPro: IPR002895; Paramesicium_SA.
 DR Pfam: PF01508; Paramesicium_SA; 21.
 DR PRINTS: PR00759; BASICPTASE.
 KW Serine protease inhibitor.
 SQ SEQUENCE 2533 AA; 264142 MW; EAED7F21E408C371 CRC64;

Query Match 85.7%; Score 36; DB 5; Length 2533;
 Best Local Similarity 57.1%; Pred. No. 46;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 I::I::I
 Db 2149 CIPITNC 2155

RESULT 7
 O90993 PRELIMINARY; PRT; 193 AA.
 AC O90993;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PP60C-SCR PROTEIN.
 GN PP60C-SCR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PECTORALIS MUSCLE;
 RX MEDLINE=90318371; PubMed=2115117;
 RA Dorai T., Wang L.-H.;
 RT "An alternative non-tyrosine protein kinase product of the c-src gene
 in chicken skeletal muscle."
 RL Mol. Cell. Biol. 10:4068-4079(1990).
 DR EMBL: M57290; AAA49078.1; -
 DR HSSP: P00523; ISRL.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 SQ SEQUENCE 193 AA; 21180 MW; 9D1B54C3B02D98 CRC64;

Query Match 83.3%; Score 35; DB 13; Length 193;
 Best Local Similarity 71.4%; Pred. No. 7.7;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 I::I::I
 Db 185 CIPITSC 191

RESULT 8
 O97887 PRELIMINARY; PRT; 389 AA.
 AC O97887;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE REISSNER'S FIBER GLYCOPROTEIN I (FRAGMENT).
 OS Bos taurus (Bovine)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SUBCOMMISSURAL ORGAN;
 RX MEDLINE=96228246; PubMed=9560467;
 RA Nualart F., Hein S., Yullis C.R., Zaraga A.M., Araya A.,
 RA Rodriguez E.M.;
 RT Cell Tissue Res. 292:239-250(1998).
 DR EMBL: AF078930; AAD17695.1; -
 DR HSSP: P56682; ICCV.
 DR InterPro: IPR000561; EGF_1like.
 DR InterPro: IPR001545; Glyco_hormone_beta.
 DR InterPro: IPR002919; TIL.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF01826; TIL; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR SMART: SM00019; EGF_1like; 1.
 DR SMART: SM00209; TSP1; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
 DR PROSITE: PS50092; TSP1; 1.
 FT NOW_TER 1
 SQ SEQUENCE 389 AA; 40720 MW; 48571E1F16F9F1B6 CRC64;

Query Match 83.3%; Score 35; DB 6; Length 389;
 Best Local Similarity 71.4%; Pred. No. 14;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 I::I::I
 Db 311 CVPVSSC 317

RESULT 9
 O9CXM8 PRELIMINARY; PRT; 462 AA.
 ID O9CXM8
 AC O9CXM8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 3110056H04RIK PROTEIN.
 GN 3110056H04RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukumishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaide I., Pong B., Quackenbush J.,
 RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 DR EMBL: AK014221; BAB29213.1; -
 DR HSSP: P00214; 1FTC.
 DR MGD: MGI:1920480; 3110056H04RLK.
 DR InterPro: IPR002919; TIL.
 DR InterPro: IPR001846; Vwd.
 DR InterPro: IPR001007; VWFC.
 DR Pfam: PF01826; TIL; 1.
 DR Pfam: PF00093; VWC; 2.
 DR Pfam: PF00094; Vwd; 1.
 DR SMART: SM00214; VWC; 2.
 DR SMART: SM00216; VWD; 1.
 DR PROSITE: PS01208; VWFC; UNKNOWN_1.
 SQ SEQUENCE 462 AA; 51641 MW; 5B6AF92A52D7A4B6 CRC64;

Query Match 83.3%; Score 35; DB 11; Length 462;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 Db 113 CVPVSSC 119

RESULT 10
 Q28161 PRELIMINARY; PRT; 727 AA.
 ID Q28161;
 AC Q28161;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PLAKOPHILIN.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 CX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95237210; PubMed=7720719;
 RA Held H.W., Schmidt A., Zimbelmann R., Schaefer S., Franke W.W.,
 RA Nuber U.A.;
 RT "Desmosomes and cytoskeletal architecture in epithelial
 RT differentiation: cell type-specific plaque components and intermediate
 RT filament anchorage.",
 RL Eur. J. Cell Biol. 65:229-245(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95196971; PubMed=7890138;
 RA Held H.W., Schmidt A., Zimbelmann R., Schaefer S.,
 RA Winter-Simanowski S., Stump S., Keith M., Fligge U., Scholzer M.,
 RA Franke W.W.;
 RT "Cell type-specific desmosomal plaque proteins of the plakoglobin
 RT family: plakophilin 1 (band 6 protein).",
 RL Differentiation 58:113-131(1994).
 DR EMBL: Z37975; CAA86029.1;
 DR InterPro: IPR000225; Armadillo.

DR Pfam: PF00514; Armadillo_seg; 5.
 DR SMART: SM00185; ARM; 4.
 DR PROSITE: PS50176; ARM_REPEAT; 3.
 SQ SEQUENCE 727 AA; 80180 MW; 3A27979279BCEDF CRC64;

Query Match 83.3%; Score 35; DB 6; Length 727;
 Best Local Similarity 85.7%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 Db 215 CVPITSC 221

RESULT 11
 Q9XSV8 PRELIMINARY; PRT; 1637 AA.
 ID Q9XSV8;
 AC Q9XSV8;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SCO-SPONDIN (FRAGMENT).
 GN SCO-SPONDIN.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 CX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gobron S., Creveaux I., Didier R., Meinel R.,
 RT "Characterization of cattle SCO-spondin.",
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A1133488; CAB46239.1; -
 DR HSSP: P56682; 1CCV.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001545; Glyco_hormone_beta.
 DR InterPro: IPR002919; TIL.
 DR InterPro: IPR000884; TSPL.
 DR InterPro: IPR001007; VWFC.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF01826; TIL; 5.
 DR Pfam: PF00090; tsp_1; 11.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR SMART: SM00068; GHB; 1.
 DR SMART: SM00209; TSPL; 11.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_2.
 DR PROSITE: PS50092; TSPL; 11.
 DR PROSITE: PS01208; VWFC; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 1637 AA; 173347 MW; 4C5BAB1D346C925 CRC64;

Query Match 83.3%; Score 35; DB 6; Length 1637;
 Best Local Similarity 71.4%; Pred. No. 51;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 Db 1206 CVPVSSC 1212

RESULT 12
 Q9VR08 PRELIMINARY; PRT; 3680 AA.
 ID Q9VR08;
 AC Q9VR08;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

01-DEC-2001 (TReMBLrel. 19, last annotation update)
 CG15637 PROTEIN.
 DP OR CG15637.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003576; AAF51000.1; -;
 DR HSSP; P35555; 1EMN.
 DR FLYBASE: FBgn0000488; dp.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR002899; EB.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR002603; ET.
 DR InterPro: IPR003645; FOLN.
 DR Pfam: PF00008; EGF_24.
 DR ProDom: PD003573; ET_1.
 DR SMART: SM00179; EGF_Ca_16.
 DR SMART: SM00001; EGF-like_35.
 DR SMART: SM00274; FOLN_5.
 DR SMART: SM00289; WRI_2.
 DR PROSITE: PS00010; ASX_HYDROXLY, 18.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 18.
 DR PROSITE: PS01187; EGF_Ca_17.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 3680 AA; 388637 MW; E5E972E1A3479EFC CRC64;

Query Match 83.3%; Score 35; DB 5; Length 3680;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVPITSC 7
 DB 588 CVPITSC 594
 RESULT 13
 Q9VJX0 PRELIMINARY; PRT; 88 AA.
 AC Q9VJX0;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE CG9260 PROTEIN.
 GN CG9260.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003640; AAF53317.1; -;
 DR FLYBASE: FBgn0023407; B4.
 SQ SEQUENCE 88 AA; 9899 MW; 27A8C83F4D0F346 CRC64;

Query Match 81.0%; Score 34; DB 5; Length 88;
 Best Local Similarity 71.4%; Pred. No. 6.3;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 1111:1
 Db 45 CVPITSC 51

RESULT 14
 P79927 PRELIMINARY; PRF: 1506 AA.

ID P79927
 AC P79927;
 DT 01-MAY-1997 (TREMUREL. 03, last sequence update)
 DT 01-MAY-1997 (TREMUREL. 03, last sequence update)
 DE 01-DEC-2001 (TREMUREL. 19, last annotation update)
 DE INTERGUMENTARY MUCIN B.1 (FRAGMENT).
 GN FIL-B.1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodidae; Xenopus.
 OC NCBI_TaxID=8353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RX MEDLINE=97153143; PubMed=8999864;
 RA Joba W., Hoffmann W.;
 RT "Similarities of intergumentary mucin B.1 (FIL-B.1) from Xenopus laevis
 and prepro-von Willebrand factor at their amino terminal regions.";
 RL J. Biol. Chem. 272:1805-1810(1997).
 DR EMBL: Y08296; CAA69604.1;
 DR Interpro: IPR000436; Sushl_SCR_CCP.
 DR Interpro: IPR002919; TIL.
 DR Interpro: IPR001846; WVD.
 DR Pfam: PF00084; Sushl; 1.
 DR Pfam: PF01826; TIL; 1.
 DR Pfam: PF00094; WVD; 3.
 DR SMART: SM00032; CCP; 1.
 DR SMART: SM00216; WVD; 3.
 FT NON_TER 1506 1506
 FT SEQUENCE 1506 AA; 163906 MW; 7BF269A74BDC817F CRC64;
 SQ

Query Match 81.0%; Score 34; DB 13; Length 1506;
 Best Local Similarity 71.4%; Pred. No. 77;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 1111:1
 Db 1292 CVPITSC 1298

RESULT 15
 ID P62635 PRELIMINARY; PRF: 1513 AA.
 AC P62635;
 DT 01-NOV-1996 (TREMUREL. 01, last sequence update)
 DT 01-NOV-1996 (TREMUREL. 01, last sequence update)
 DE 01-DEC-2001 (TREMUREL. 19, last annotation update)
 DE MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2) (FRAGMENT).
 GN MUC2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RX MEDLINE=94299489; PubMed=8027037;
 RA Ohmori H., Bohman A.F., Gallup M., Tsuda T., Kai H., Gum J.R. Jr.,
 Kim Y.S., Basbaum C.B.;
 RT "Molecular cloning of the amino-terminal region of a rat MUC 2 mucin
 gene homologue. Evidence for expression in both intestine and
 airway.";
 RL J. Biol. Chem. 269:17833-17840(1994).
 SQ

RN [2]
 RP SEQUENCE OF 21-36.
 RC TISSUE=INTESTINAL EPITHELIUM;
 RX MEDLINE=98180965; PubMed=9512496;
 RA Khatri I.A., Forstner G.G., Forstner J.F.;
 RT "Susceptibility of the cysteine-rich N-terminal and C-terminal ends of
 rat intestinal mucin Muc 2 to proteolytic cleavage.";
 RL Biochem. J. 331:323-330(1998).
 CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
 OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
 CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
 CC AGENTS AT MUCOSAL SURFACES.
 CC -1- SUBUNIT: MULTIMERIC.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: EXPRESSION IN INTESTINE AND AIRWAY.
 DR EMBL: U07615; AAA21655.2;
 DR Interpro: IPR002965; P-rich_extensn.
 DR Interpro: IPR002919; TIL.
 DR Interpro: IPR001846; WVD.
 DR Interpro: IPR001007; WVEC.
 DR Pfam: PF01826; TIL; 2.
 DR Pfam: PF00094; WVD; 3.
 DR PRINTS: PR01217; PRICHEXTENS.
 DR SMART: SM00214; WVC; 1.
 DR SMART: SM00216; WVD; 3.
 DR Signal; Repeat; Glycoprotein.
 FT CHAIN 1 20
 FT DOMAIN 21 1513
 FT REPEAT 1392 1407
 FT REPEAT 1408 1423
 FT REPEAT 1424 1434
 FT REPEAT 1435 1445
 FT REPEAT 1446 1456
 FT REPEAT 1457 1467
 FT REPEAT 1468 1478
 FT REPEAT 1479 1489
 FT REPEAT 1490 1500
 FT REPEAT 1501 1511
 FT REPEAT 1512 1513
 FT REPEAT 1512 1513
 FT CARBOHYD 160 160
 FT CARBOHYD 420 420
 FT CARBOHYD 667 667
 FT CARBOHYD 767 767
 FT CARBOHYD 837 837
 FT CARBOHYD 892 892
 FT CARBOHYD 1136 1136
 FT CARBOHYD 1151 1151
 FT CARBOHYD 1212 1212
 FT CARBOHYD 1227 1227
 FT CARBOHYD 1243 1243
 FT CARBOHYD 1350 1350
 FT NON_TER 1513 1513
 FT SEQUENCE 1513 AA; 166036 MW; 26109DCA1BE7D008 CRC64;
 SQ

Query Match 81.0%; Score 34; DB 11; Length 1513;
 Best Local Similarity 71.4%; Pred. No. 77;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPITSC 7
 1111:1
 Db 709 CVPITSC 715

Search completed: June 17, 2002, 16:23:42
 Job time: 1616 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:38 ; Search time 75.98 Seconds
(Without alignments)
2.250 Million cell updates/sec

Title: US-09-761-636A-11
Perfect score: 42
Sequence: 1 CVPITSC 7

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/prodata/2/iaa/5A.COMB.pep:*
- 2: /cgn2_6/prodata/2/iaa/5B.COMB.pep:*
- 3: /cgn2_6/prodata/2/iaa/6A.COMB.pep:*
- 4: /cgn2_6/prodata/2/iaa/6B.COMB.pep:*
- 5: /cgn2_6/prodata/2/iaa/PTUS.COMB.pep:*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	81.0	2476	2	US-08-276-967-2
2	34	81.0	2594	4	US-08-718-388-7
3	34	81.0	5405	4	US-08-718-388-9
4	33	78.6	280	2	US-08-969-106-11
5	33	78.6	376	2	US-08-875-972-2
6	32	76.2	661	2	US-08-417-174-27
7	32	76.2	661	2	US-08-417-174-121
8	32	76.2	661	2	US-08-231-565A-27
9	32	76.2	661	2	US-09-007-961-27
10	32	76.2	661	4	US-09-267-439-27
11	32	76.2	661	4	US-09-267-439-121
12	32	76.2	668	1	US-07-891-942G-6
13	31	73.8	7	1	US-08-526-710-21
14	31	73.8	7	3	US-08-862-855-21
15	31	73.8	7	4	US-09-226-885-21
16	31	73.8	7	4	US-09-227-906-21
17	31	73.8	12	3	US-08-742-243-35
18	31	73.8	12	3	US-08-742-243-36
19	31	73.8	12	3	US-08-742-243-37
20	31	73.8	12	3	US-08-742-243-38
21	31	73.8	17	3	US-08-807-992B-23
22	31	73.8	109	3	US-08-691-794-3
23	31	73.8	110	4	US-09-392-932-11
24	31	73.8	120	6	5194596-9
25	31	73.8	120	6	5219739-9
26	31	73.8	121	6	5194596-19
27	31	73.8	121	6	5219739-20

28	31	73.8	145	3	US-08-784-551C-2	Sequence 2, Appl
29	31	73.8	145	4	US-09-392-932-2	Sequence 2, Appl
30	31	73.8	146	4	US-08-586-039B-33	Sequence 3, Appl
31	31	73.8	147	3	US-08-807-992B-1	Sequence 1, Appl
32	31	73.8	147	4	US-09-392-932-1	Sequence 1, Appl
33	31	73.8	164	6	5194596-17	Patent No. 5194596
34	31	73.8	164	6	5219739-17	Patent No. 5219739
35	31	73.8	164	6	5219739-18	Patent No. 5219739
36	31	73.8	165	6	5194596-18	Patent No. 5194596
37	31	73.8	165	6	5219739-19	Patent No. 5219739
38	31	73.8	189	1	US-08-469-427A-15	Sequence 15, Appl
39	31	73.8	190	2	US-08-569-063C-20	Sequence 20, Appl
40	31	73.8	190	4	US-08-586-039B-31	Sequence 31, Appl
41	31	73.8	190	6	5332671-3	Patent No. 5332671
42	31	73.8	191	3	US-08-567-200A-2	Sequence 2, Appl
43	31	73.8	191	3	US-08-807-992B-2	Sequence 2, Appl
44	31	73.8	191	3	US-08-691-794-2	Sequence 2, Appl
45	31	73.8	191	4	US-08-795-430-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5851817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; TITLE OF INVENTION: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,967
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-967-2

Query Match 81.0%; Score 34; DB 2; Length 2476;
Best Local Similarity 71.4%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CVPITSC 7
||||:|

Db 1505 CVELSQC 1511

RESULT 2

US-08-718-388-7

Sequence 7, Application US/08718388
Patent No. 6271362

GENERAL INFORMATION:

APPLICANT: MORIKAWA, MINORU

APPLICANT: MORIKAWA, MINORU

TITLE OF INVENTION: GENE ENCODING IGG FC REGION-BINDING

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

STREET: PO BOX 747

CITY: FALLS CHURCH

STATE: VA

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/718,388

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: MURPHY JR, GERALD M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 0230-111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2594 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 81.0%; Score 34; DB 4; Length 2594;

Best local Similarity 71.4%; Pred. No. 5.9e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVELTSC 7

Db 1183 CLEPLASC 1189

RESULT 3

US-08-718-388-9

Sequence 9, Application US/08718388
Patent No. 6271362

GENERAL INFORMATION:

APPLICANT: MORIKAWA, MINORU

APPLICANT: MORIKAWA, MINORU

TITLE OF INVENTION: GENE ENCODING IGG FC REGION-BINDING

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

STREET: PO BOX 747

CITY: FALLS CHURCH

STATE: VA

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/718,388

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: MURPHY JR, GERALD M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 0230-111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 5405 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match

Best local Similarity 81.0%; Score 34; DB 4; Length 5405;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVELTSC 7

Db 1183 CLEPLASC 1189

RESULT 4

US-08-969-106-11

Sequence 11, Application US/08969106
Patent No. 5986055

GENERAL INFORMATION:

APPLICANT: Yang, M.

APPLICANT: Nandabalan, K.

APPLICANT: Schultz, V.

TITLE OF INVENTION: CDK2 INTERACTIONS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/969,106

FILING DATE: 13-NOV-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mastro, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7934-057

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 280 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-969-106-11

Query Match 78.6%; Score 33; DB 2; Length 280;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CVPLTSC 7
111111
DB 227 CVPATEC 233

RESULT 5

US-08-875-972-2
; Sequence 2, Application US/08875972
; Patent No. 5985564
; GENERAL INFORMATION:
; APPLICANT: Huntington Potter and Jinhue Li
; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
; TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,972
; FILING DATE: 08-AUG-97
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,448
; FILING DATE: 16-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan Esq., Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-03PA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-875-972-2

Query Match 78.6%; Score 33; DB 2; Length 376;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPLTSC 7
111111
DB 26 CVPATEC 32

RESULT 6

US-08-417-174-27
; Sequence 27, Application US/08417174
; Patent No. 5844075
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,174
; FILING DATE: 05-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,565
; FILING DATE: 22-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CAROL M. GRUPPI
; REGISTRATION NUMBER: 37,341
; REFERENCE/DOCKET NUMBER: 2026-4124051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: protein
; US-08-417-174-27

Query Match 76.2%; Score 32; DB 2; Length 661;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPLTSC 7
111111
DB 296 IPLTSC 301

RESULT 7

US-08-417-174-121
; Sequence 121, Application US/08417174
; Patent No. 5844075
; GENERAL INFORMATION:
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
; APPLICANT: STEVEN A.
; TITLE OF INVENTION: MELANOMA ANTIGENS AND
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 126
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/417.174
FILING DATE: 05-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/231.565
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-08-417-174-121

Query Match 76.2%; Score 32; DB 2; Length 661;
Best Local Similarly 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 VPLTSC 7
DB 296 IPLTSC 301

RESULT 8
US-08-231-565A-27
Sequence 27, Application US/08231565A
Patent No. 5874560
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231.565A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown

TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-08-231-565A-27

Query Match 76.2%; Score 32; DB 2; Length 661;
Best Local Similarly 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 VPLTSC 7
DB 296 IPLTSC 301

RESULT 9
US-09-007-961-27
Sequence 27, Application US/09007961
Patent No. 5994523
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/007,961
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/231.565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-09-007-961-27

Query Match 76.2%; Score 32; DB 2; Length 661;
Best Local Similarly 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPLTSC 7
DB 296 IPLTSC 301

RESULT 10
US-09-267-439-27

Sequence 27, Application US/09267439
Patent No. 6270778
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,439
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
APPLICATION NUMBER: US/08/231,565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-09-267-439-27

Query Match 76.2%; Score 32; DB 4; Length 661;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPLTSC 7
Db 296 IPLTSC 301

RESULT 11
US-09-267-439-121
Sequence 121, Application US/09267439
Patent No. 6270778
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA

ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,439
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
APPLICATION NUMBER: US/08/231,565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-09-267-439-121

Query Match 76.2%; Score 32; DB 4; Length 661;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPLTSC 7
Db 296 IPLTSC 301

RESULT 12
US-07-891-942G-6
Sequence 6, Application US/07891942G
Patent No. 5679511
GENERAL INFORMATION:
APPLICANT: Kwon, Byoung Se
TITLE OF INVENTION: CDNA CLONES FOR HUMAN TYROSINASE AND FOR
A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christopher A. Michaels, Barnard, Brown &
STREET: 306 East State Street, Suite 220
CITY: Ithaca
STATE: NY
COUNTRY: USA
ZIP: 14850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/891,942G
FILING DATE: 01-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/915,753
FILING DATE: 06-OCT-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/362,847
FILING DATE: 07-JUN-1989

ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A
REGISTRATION NUMBER: 34,390
REFERENCE/DOCKET NUMBER: IND1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 607-273-1711
TELEFAX: 607-273-2609
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 668 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PUBLICATION INFORMATION:
AUTHORS: Kwon, Byoung Se
AUTHORS: Chintamani, Chaya
AUTHORS: Kozak, Christine A
AUTHORS: Copeland, Neal G
AUTHORS: Gilbert, Debra J
AUTHORS: Jenkins, Nancy
AUTHORS: Barton, David
AUTHORS: Francke, Uta
AUTHORS: Kobayashi, Yvonne
AUTHORS: Kim, Kack K
TITLE: A melanocyte-specific gene, Pmel 17, maps
TITLE: near the silver coat color locus on mouse
TITLE: chromosome 10 and is in a syntenic region on human
TITLE: chromosome 12
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 9228-9232
DATE: October-1991
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 668
US-07-891-942G-6

Query Match 76.2%; Score 32; DB 1; Length 668;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPLTSC 7
Db 296 IPLTSC 301

RESULT 13
US-08-526-710-21
Sequence 21, Application US/08526710
Patent No. 5622699
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki
APPLICANT: Pasqualini, Renata
TITLE OF INVENTION: Method of identifying molecules that
TITLE OF INVENTION: Home to a selected organ in vivo
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,710
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1779
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-526-710-21

Query Match 73.8%; Score 31; DB 1; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTSC 7
Db 1 CLPVASC 7

RESULT 14
US-08-862-855-21
Sequence 21, Application US/08862855
Patent No. 6068829
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki
APPLICANT: Pasqualini, Renata
TITLE OF INVENTION: Method of identifying molecules that
TITLE OF INVENTION: Home to a selected organ in vivo
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/862,855
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,710
FILING DATE: 11-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/813,273
FILING DATE: 10-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2621
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-862-855-21

Query Match 73.8%; Score 31; DB 3; Length 7;

Best Local Similarity 57.1%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPPTSC 7
1:1:11
Db 1 CLPVASC 7

RESULT 15

US-09-226-985-21
; Sequence 21, Application US/092226985
; Patent No. 6296832

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/526,710

; FILING DATE: 11-SEP-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/813,273

; FILING DATE: 10-MAR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/862,855

; FILING DATE: 23-MAY-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-LJ 3423

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-226-985-21

Query Match

Best Local Similarity 73.8%; Score 31; DB 4; Length 7;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPPTSC 7
1:1:11
Db 1 CLPVASC 7

Search completed: June 17, 2002, 15:56:39
Job time: 88 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:25 ; Search time 95.43 Seconds
(without alignments)
6.041 Million cell updates/sec

Title: US-09-761-636a-12

Perfect score: 38

Sequence: 1 CVPLTC 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR.71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	817	2 T25674	hypothetical prote
2	37	97.4	293	2 T27621	hypothetical prote
3	35	92.1	1627	2 S65464	pregnancy-associat
4	34	89.5	367	2 T20914	hypothetical prote
5	33	86.8	332	2 T48933	adenosine receptor
6	33	86.8	415	2 A54176	endothelin recepto
7	33	86.8	467	2 AC2390	transposase all467
8	33	86.8	637	2 T08050	probable ethylene
9	33	86.8	650	2 A34498	glycoprotein anti
10	33	86.8	1236	2 A53970	1-phosphatidylinos
11	32	84.2	108	2 S12982	hypothetical 12.2K
12	32	84.2	137	2 T26433	hypothetical prote
13	32	84.2	156	2 C81364	probable periplasm
14	32	84.2	252	2 B82659	two-component syst
15	32	84.2	301	2 T40593	cytoplasmic dynein
16	32	84.2	314	2 B43652	ferredoxin [2Fe-2S
17	32	84.2	345	1 MBRO	apolipoprotein H p
18	32	84.2	398	2 A35281	integumentary muc
19	32	84.2	425	2 T40610	probable clathrin
20	32	84.2	518	2 AD1930	ammonium transport
21	32	84.2	972	2 A30363	glycoprotein GP330
22	32	84.2	1034	2 JC5598	mucin - rat
23	32	84.2	4660	2 T42737	gp330 protein prec
24	31	81.6	40	2 E45495	beta-defensin-5 -
25	31	81.6	133	2 T29606	hypothetical prote
26	31	81.6	224	2 T27063	hypothetical prote
27	31	81.6	326	2 C69341	cobalt transport A
28	31	81.6	414	2 T50019	hypothetical prote
29	31	81.6	434	2 I50132	vasotocin receptor

30	31	81.6	471	2 AC3622	mannose-6-phosphat
31	31	81.6	588	2 T24121	hypothetical prote
32	31	81.6	718	2 JE0188	RES4-25 protein -
33	31	81.6	822	2 T01622	probable salt-indu
34	31	81.6	869	1 JC4858	VDL receptor prec
35	31	81.6	1042	2 T26644	hypothetical prote
36	31	81.6	1313	2 T29027	hypothetical prote
37	31	81.6	1451	2 T30821	P-glycoprotein - S
38	31	81.6	2098	2 T13166	rough deal protein
39	31	81.6	2946	2 T00867	hypothetical prote
40	30	78.9	40	2 G45495	beta-defensin-7 -
41	30	78.9	40	2 I45495	beta-defensin-9 -
42	30	78.9	42	2 C45495	beta-defensin-3 -
43	30	78.9	42	2 F45495	beta-defensin-6 -
44	30	78.9	64	2 A47438	airway epithelial
45	30	78.9	64	2 A56128	lingual antimicrob

ALIGNMENTS

RESULT 1
T25674
hypothetical protein F08D12.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T25674
R/Ref: T. Waterston, R.
Submitted to the EMBL Data Library, December 1996
A/Description: The sequence of C. elegans cosmid F08D12.
A/Reference number: Z20068
A/Accession: T25674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-817 <LEF>
A/Cross-references: EMBL:U08040; PIDN:AA37929.1; GSPDB:GN00020; CESP:F08D12.6
A:Experimental source: strain Bristol N2; clone F08D12
C/Genetics:
A:Gene: CESP:F08D12.6
A:Map position: 2
A:Introns: 27/2; 329/3; 407/1; 477/3; 538/3; 622/2

Query Match 100.0%; Score 38; DB 2; Length 817;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CVPLTC 6
Db 479 CVPLTC 484
RESULT 2
T27621
hypothetical protein ZC504.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T27621
R/Kershaw, J.
Submitted to the EMBL Data Library, July 1995
A/Reference number: Z20394
A/Accession: T27621
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-293 <WIL>
A/Cross-references: EMBL:Z50029; PIDN:CAA90343.1; GSPDB:GN00028; CESP:ZC504.5
A:Experimental source: clone ZC504
C/Genetics:
A:Gene: CESP:ZC504.5
A:Map position: X
A:Introns: 40/3; 140/3; 173/3; 255/3

Query Match 97.4%; Score 37; DB 2; Length 293;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
|:|:|:|
Db 123 CIPVTC 128

RESULT 3

S65464
pregnancy-associated plasma protein A precursor - human
N:Alternate names: PAPP-A
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 05-Nov-1999
C:Accession: S65464; A54220; I38097
R:Haanling, J.; Oxvig, C.; Overgaard, M.T.; Ebbesen, P.; Kristensen, T.; Sotttrup-Jensen, S.
submited to the EMBL Data Library, June 1995
A:Description: Complete cDNA sequence of the preproform of human pregnancy-associated plasma protein A.
A:Reference number: S65464
A:Accession: S65464
A:Molecule type: mRNA

A:Residues: 1-1627 <RHA>
A:Cross-references: EMBL:U28727; NID:g1142969; PIDN:AAC50543.1; PID:g1142970
R:Haanling, J.; Oxvig, C.; Overgaard, M.T.; Ebbesen, P.; Kristensen, T.; Sotttrup-Jensen, S.
Eur. J. Biochem. 237, 159-163, 1996
A:Title: Complete cDNA sequence of the preproform of human pregnancy-associated plasma protein A.
A:Reference number: S65463; MUID:96203921
A:Accession: S65463
A:Molecule type: mRNA

A:Residues: 1-102 <HAW>
A:Cross-references: EMBL:U28727
R:Kristensen, T.; Oxvig, C.; Sand, O.; Moller, N.P.H.; Sotttrup-Jensen, L.
Biochemistry 33, 1592-1598, 1994
A:Title: Amino acid sequence of human pregnancy-associated plasma protein-A derived from placenta.
A:Reference number: A54220; MUID:94146014
A:Accession: A54220
A:Molecule type: mRNA

A:Residues: 77-1627 <RES>
A:Cross-references: GB:X68280; NID:g394649; PIDN:CAA48341.1; PID:g394650
R:Oxvig, C.; Sand, O.; Kristensen, T.; Glatch, G.J.; Sotttrup-Jensen, L.
J. Biol. Chem. 268, 12243-12246, 1993
A:Title: Circulating human pregnancy-associated plasma protein-A is disulfide-bridged to placental alpha-2-macroglobulin.
A:Reference number: I38097; MUID:93286045
A:Accession: I38097
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 77-1627 <RES>
A:Cross-references: EMBL:X68280; NID:g394649; PIDN:CAA48341.1; PID:g394650
C:Genetics:
A:Gene: GDB:PAPPA
A:Cross-references: GDB:134729; OMIM:176385
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-80/Domain: propeptide #status predicted <PRO>
F:81-1627/Product: pregnancy-associated plasma protein A #status predicted <MAT>

Query Match 92.1%; Score 35; DB 2; Length 1627;
Best Local Similarity 83.3%; Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
|:|:|:|
Db 1410 CVPVTC 1415

RESULT 4

T20914
hypothetical protein F14F8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T20914
R:Liroyd, C.
Submitted to the EMBL Data Library, March 1997
A:Reference number: Z19346
A:Accession: T20914
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-367 <WIL>
A:Cross-references: EMBL:Z92782; PIDN:CA807186.1; GSPDB:GN00023; CESP:F14F8.5
A:Experimental source: clone F14F8
C:Genetics:
A:Gene: CESP:F14F8.5
A:Map position: 5
A:Introns: 23/3; 97/2; 145/2; 301/2

Query Match 89.5%; Score 34; DB 2; Length 367;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
|:|:|:|
Db 92 CIPVTC 97

RESULT 5

I48933
adenosine receptor subtype - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: I48933
R:Margardt, D.L.; Walker, L.L.; Heinemann, S.
J. Immunol. 152, 4508-4515, 1994
A:Title: Cloning of Two Adenosine Receptor Subtypes From Mouse Bone Marrow-Derived Cells.
A:Reference number: I48931; MUID:94205670
A:Accession: I48933
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-332 <RES>
A:Cross-references: EMBL:U05673; NID:g498337; PIDN:AAA19001.1; PID:g498338
C:Superfamily: adenosine receptor A1

Query Match 86.8%; Score 33; DB 2; Length 332;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPLTC 6
|:|:|:|
Db 166 CIPVTC 171

RESULT 6

A54126
endothelin receptor-AX - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 16-Jul-1999
C:Accession: A54126
R:Kumar, C.; Mwangi, V.; Nuthuaganti, P.; Wu, H.L.; Pullen, M.; Brun, K.; Alyar, I.
J. Biol. Chem. 269, 13414-13420, 1994
A:Title: Cloning and characterization of a novel endothelin receptor from Xenopus laevis.
A:Reference number: A54126; MUID:94230448
A:Accession: A54126
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-415 <KUM>
A:Cross-references: GB:U06633; NID:g507893; PIDN:AAA19570.1; PID:g507894
A:Note: authors translated the codon ATG for residue 84 as His
C:Superfamily: endothelin receptor B

Query Match 86.8%; Score 33; DB 2; Length 415;
Best Local Similarity 83.3%; Pred. No. 86;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||||
Db 254 CVPLAC 259

RESULT 7
AC2390
Transposase all4675 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AC2390

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MID:21595285; PMID:11759840

A:Accession: AC2390

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-467 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076374.1; PID:g17133812; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4675

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 467;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||||
Db 61 CCPLTC 66

RESULT 8

T08050
probable ethylene receptor - muskmelon

C:Species: Cucumis melo (muskmelon)

C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 08-Oct-1999

C:Accession: T08050

R:Sato-Nara, K.; Yuhashi, K.; Higashi, K.; Hosoya, K.; Kubota, M.; Ezura, H.

Submitted to the EMBL Data Library, December 1997

A:Description: Expression of an ERK homolog gene of melon (Cucumis melo L. reticulatus)

A:Reference number: Z16317

A:Accession: T08050

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-637 <SMT>

A:Cross-references: EMBL:AF037368; NID:g2707333; PIDN:AA099477.1; PID:g2707334

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 637;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||||
Db 226 CVPYTC 231

RESULT 9

A34498
glycoprotein antigen Bm86 precursor - southern cattle tick

C:Species: Boophilus microplus (southern cattle tick)

C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 11-Jan-2000

C:Accession: A34498

R:Rand, K.N.; Moore, T.; Srisankha, A.; Spring, K.; Tellam, R.; Willadsen, P.; Cobon, G.

Proc. Natl. Acad. Sci. U.S.A. 86, 9657-9661, 1989

A:Title: Cloning and expression of a protective antigen from the cattle tick Boophilus

A:Reference number: A34498; MID:30099323

A:Accession: A34498

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-650 <RAN>

A:Cross-references: GB:M29321; NID:g161667; PID:g161668

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: glycoprotein

F:209-246/Domain: EGF homology <EGF>

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 650;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||||
Db 147 CVPYTC 152

RESULT 10

A53970
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) gamma-D -

C:Species: Drosophila melanogaster

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000

C:Accession: A53970

R:Emori, Y.; Sugaya, R.; Akimaru, H.; Higashijima, S.; Shishido, E.; Saigo, K.; Ho

J. Biol. Chem. 269, 19474-19479, 1994

A:Title: Drosophila phospholipase C-gamma expressed predominantly in blastoderm ce

A:Reference number: A53970; MID:94308233

A:Accession: A53970

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1236 <EMO>

A:Cross-references: GB:D29806

C:Genetics:

A:Gene: plc-gamma4

A:Cross-references: FlyBase:FBgn0003416

C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase II; 1-ph

hodiesterase domain Y homology; SH2 homology; SH3 homology

C:Keywords: phosphoric diester hydrolase

F:326-468/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain

F:569-668/Domain: SH2 homology <SH2>

F:701-789/Domain: SH2 homology <SH2>

F:831-879/Domain: SH3 homology <SH3>

F:979-1095/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase doma

Query Match
Best Local Similarity 86.8%; Score 33; DB 1; Length 1236;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||||
Db 1196 CVPLTC 1201

RESULT 11
SI2982
hypothetical 12.2k protein (68k signal recognition particle region) - dog

C:Species: Canis lupus familiaris (dog)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999

C:Accession: SI2982

R:Herz, J.; Flint, N.; Stanley, K.; Frank, R.; Dobberstein, B.

FEBS Lett. 276, 103-107, 1990

A:Title: The 68 kDa protein of signal recognition particle contains a glycine-rich

A:Reference number: SI2981; MID:91092392

A:Accession: SI2982

A:Molecule type: mRNA

A:Residues: 1-108 <HER>

A:Cross-references: GB:X53744; NID:g931; PIDN:CAA37774.1; PID:g933

C:Superfamily: dog hypothetical 12.2k protein (68k signal recognition particle re

Query Match	84.2%	Score 32;	DB 2;	Length 108;
Best Local Similarity	66.7%	Pred. NO. 43;		
Matches	4;	Conservative	2;	Mismatches
			0;	Indels
				0;
Gaps				0;
Qy	1	CVP/LNC	6	
Db	20	CVP/MSC	25	

Mon Jun 17 16:11:10 2002

us-09-761-636a-12.open.rpr

Page 5

Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:30 ; Search time 44.84 Seconds
(without alignments)
5.181 Million cell updates/sec

Title: US-09-761-636A-12
Perfect score: 38
Sequence: 1 CVPLTC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	35	92.1	530	1	PDP2_RAT
2	35	92.1	1627	1	PAPA_HUMAN
3	33	86.8	332	1	AA2B_MOUSE
4	33	86.8	650	1	BM86_BOOMI
5	33	86.8	1587	1	LMG3_HUMAN
6	32	84.2	314	1	VANB_PSES9
7	32	84.2	345	1	APOH_BOVIN
8	32	84.2	370	1	DCAM_TRYBB
9	32	84.2	398	1	MOBI_XENLA
10	32	84.2	940	1	CHRD_BRARE
11	32	84.2	1581	1	LMG3_MOUSE
12	32	84.2	4655	1	LRP2_HUMAN
13	32	84.2	4660	1	LRP2_RAT
14	31	81.6	64	1	BD05_BOVIN
15	31	81.6	167	1	OB_SMICR
16	31	81.6	188	1	VO60_FOMFV
17	31	81.6	434	1	AVT_CATCO
18	31	81.6	810	1	NEIL_HUMAN
19	30	78.9	38	1	BD08_BOVIN
20	30	78.9	40	1	BD02_BOVIN
21	30	78.9	40	1	BD07_BOVIN
22	30	78.9	42	1	BD06_BOVIN
23	30	78.9	55	1	BD09_BOVIN
24	30	78.9	57	1	BD03_BOVIN
25	30	78.9	64	1	EAR_BOVIN
26	30	78.9	64	1	LAP_BOVIN
27	30	78.9	64	1	TAP_BOVIN
28	30	78.9	194	1	KRUB_HUMAN
29	30	78.9	241	1	BTR_BOPE
30	30	78.9	261	1	YK83_CAEEL
31	30	78.9	270	1	ETLR_MOUSE
32	30	78.9	426	1	ETLR_RAT
33	30	78.9	427	1	ETLR_BOVIN

34	30	78.9	427	1	ETLR_HUMAN	P25101 homo sapien
35	30	78.9	427	1	ETLR_PIG	029010 sus scrofa
36	30	78.9	444	1	ETLR_XENLA	P32940 xenopus lae
37	30	78.9	445	1	MDM2_BRARE	042354 brachydanio
38	30	78.9	454	1	ATRY_HUMAN	P17735 homo sapien
39	30	78.9	454	1	ATRY_RAT	P04694 rattus norv
40	30	78.9	596	1	AGP2_YEAST	P38090 saccharomyc
41	30	78.9	781	1	K6PE_SCHMA	027778 schistosoma
42	30	78.9	810	1	TLAR_MOUSE	P16382 mus musculu
43	30	78.9	1286	1	CTRO_HUMAN	014578 homo sapien
44	30	78.9	1597	1	CTRO_MOUSE	P49025 mus musculu
45	29	76.3	38	1	BD11_BOVIN	P46169 bos taurus

ALIGNMENTS

```

RESULT 1
PDP2_RAT ID STANDARD: PRT: 530 AA.
AC 088484:
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 2, mitochondrial
DE precursor (EC 3.1.3.43) (PDP 2) (Pyruvate dehydrogenase phosphatase,
DE catalytic subunit 2) (PP2C 2).
GN PDP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=98316337; PubMed=9651365;
RX Huang B., Gudi R., Wu P., Harris R.A., Hamilton J., Popov K.M.;
RT "Isoenzymes of pyruvate dehydrogenase phosphatase. DNA-derived amino
RT acid sequences, expression, and regulation.";
RL J. Biol. Chem. 273:17680-17688(1998).
CC -1- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
CC REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE
CC PYRUVATE DEHYDROGENASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC -1- COFACTOR: MAGNESIUM.
CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A FAD PROTEIN OF
CC UNKNOWN FUNCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER.
CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF062741; AAC40168.1; -.
CC InterPro: IPR000222; PP2C.
CC InterPro: IPR001932; PP2C_domain.
CC Pfam: PF00481; PP2C_2.
CC SMART: SM00331; PP2C_Sig; 1.
CC SMART: SM00332; PP2C_C; 1.
CC PROSITE: PS01032; PP2C; 1.
CC HYDROLASE; Mitochondrion; Transist peptide; Magnesium.
CC TRANSIT 1 67 MITOCHONDRION (POTENTIAL).
CC CHAIN 68 530 [PYRUVATE DEHYDROGENASE (LIPONAMIDE)]-
CC PHOSPHATASE 2.
CC SEQUENCE 530 AA; 59654 MW; 5AB8688FAC7BAD9CD CRC64;

```

Query Match 92.1%; Score 35; DB 1; Length 530;
Best Local Similarity 83.3%; Pred. No. 8.9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVP/LTC 6
1:1111
DB 303 CLEP/LTC 308

RESULT 2
PAPA_HUMAN STANDARD: PRT; 1627 AA.
ID PAPA_HUMAN Q08371; Q9UDK7;
AC Q13219; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Pregnancy-associated plasma protein-A precursor (EC 3.4.24.-) (PAPP-A)
DE (insulin-like growth factor-dependent IGF binding protein-4 protease)
DE (IGF-dependent IGFBP-4 protease) (IGFBP-4ase).
GN PAPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RP TISSUE=Placenta;
RX MEDLINE=96203921; PubMed=8620868;
RA Haaning J., Oxvig C., Overgaard M.T., Ebbesen P., Kristensen T.,
RT Sottrup-Jensen L.,
RT "Complete cDNA sequence of the preproform of human pregnancy-
RT associated plasma protein-A. Evidence for expression in the brain and
RT induction by CAMP."
RL Eur. J. Biochem. 237:159-163(1996).
RN [2]
RP SEQUENCE OF 77-1627 FROM N.A., SEQUENCE OF 81-98; 117-126; 210-224;
RP 466-485; 507-519; 576-593; 609-621; 718-736; 742-754; 1006-1017;
RP 1259-1273; 1369-1374; 1389-1398; 1490-1509; 1524-1533 AND 1537-1544,
RP VARIANT SER-944, AND TISSUE SPECIFICITY.
RP TISSUE=Placenta, AND Serum;
RX MEDLINE=94146014; PubMed=7508748;
RA Kristensen T., Oxvig C., Sand O., Møller N.P.H., Sottrup-Jensen L.,
RT "Amino acid sequence of human pregnancy-associated plasma protein-A
RT derived from cloned cDNA."
RL Biochemistry 33:1592-1598(1994).
RN [3]
RP SEQUENCE OF 81-89; 117-126; 210-224; 460-485; 507-519; 576-593;
RP 718-736; 742-754; 1259-1273; 1369-1374; 1490-1509; 1524-1533 AND
RP 1537-1544, SUBUNITS, AND INTERCHAIN DISULFIDE BOND.
RP TISSUE=Serum;
RX MEDLINE=93286045; PubMed=7685339;
RA Oxvig C., Sand O., Kristensen T., Gleich G.J., Sottrup-Jensen L.,
RT "Circulating human pregnancy-associated plasma protein-A is disulfide-
RT bridged to the proform of eosinophil major basic protein."
RL J. Biol. Chem. 268:12243-12246(1993).
RN [4]
RP IDENTIFICATION, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP SPECIFICITY.
RP TISSUE=Placenta;
RX MEDLINE=99179030; PubMed=10077652;
RA Lawrence J.B., Oxvig C., Overgaard M.T., Sottrup-Jensen L.,
RT Gleich G.J., Hays L.G., Yates J.R. III, Conover C.A.,
RT "The insulin-like growth factor (IGF)-dependent IGF binding protein-4
RT protease secreted by human fibroblasts is pregnancy-associated plasma
RT protein-A."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3149-3153(1999).
RN [5]
RP FUNCTION, SUBUNITS, AND ENZYME REGULATION.
RX MEDLINE=20469470; PubMed=10913121;
RA Overgaard M.T., Haaning J., Boldt H.B., Olsen I.M., Laursen L.S.,
RA Christensen M., Gleich G.J., Sottrup-Jensen L., Conover C.A.,
RA Oxvig C.,

RT "Expression of recombinant human pregnancy-associated plasma protein-A
RT and identification of the proform of eosinophil major basic protein
RT as its physiological inhibitor."
RL J. Biol. Chem. 275:31128-31133(2000).
RN [6]
RP TISSUE SPECIFICITY.
RX MEDLINE=95057018; PubMed=7526035;
RA Bonno M., Oxvig C., Kephart G.M., Wagner J.M., Kristensen T.,
RA Sottrup-Jensen L., Gleich G.J.,
RT "Localization of pregnancy-associated plasma protein-A and
RT colocalization of pregnancy-associated plasma protein-A messenger
RT ribonucleic acid and eosinophil granule major basic protein messenger
RT ribonucleic acid in placenta."
RL Lab. Invest. 71:560-566(1994).
RN [7]
RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX MEDLINE=99423540; PubMed=10491647;
RA Overgaard M.T., Oxvig C., Christensen M., Lawrence J.B.,
RA Conover C.A., Gleich G.J., Sottrup-Jensen L., Haaning J.,
RT "Messenger ribonucleic acid levels of pregnancy-associated plasma
RT protein-A and the proform of eosinophil major basic protein:
RT expression in human reproductive and nonreproductive tissues."
RL Biol. Reprod. 61:1083-1089(1999).
RN [8]
RP DEVELOPMENTAL STAGE.
RX MEDLINE=95293954; PubMed=7539791;
RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
RA Stigbrand T., Gleich G.J., Sottrup-Jensen L.,
RT "Identification of angiotensinogen and complement C3dg as novel
RT proteins binding the proform of eosinophil major basic protein in
RT human pregnancy serum and plasma."
RL J. Biol. Chem. 270:13645-13651(1995).
CC -1- FUNCTION: Metalloprotease which specifically cleaves IGFBP-4 in
CC the presence of IGF, resulting in release of bound IGF.
CC -1- ENZYME REGULATION: Inhibited by complexation with the proform
CC of PRG2.
CC -1- SUBUNIT: Homodimer; disulfide-linked. In pregnancy serum,
CC predominantly found as a disulfide-linked 2:2 heterotetramer with
CC the proform of PRG2.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: High levels in placenta and pregnancy serum.
CC In placenta, expressed in X cells in septa and anchoring villi,
CC and in syncytiotrophoblasts in the chorionic villi. Lower levels
CC are found in a variety of other tissues including kidney,
CC myometrium, endometrium, ovaries, breast, prostate, bone marrow,
CC colon, fibroblasts and osteoblasts.
CC -1- DEVELOPMENTAL STAGE: Present in serum and placenta during
CC pregnancy; levels increase throughout pregnancy.
CC -1- INDUCTION: By 8-bromodeoxyadenosine-3',5'-phosphate.
CC -1- PTM: There appear to be no free cysteinyl groups.
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M46.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).
CC
CC EMBL: U28727; AAC50543.1; -;
CC EMBL: X68280; CAA48341.1; -;
CC MIM: 176385; -;
CC MEROPS: M46.001; -;
CC InterPro: IPR000800; Notch.
CC InterPro: IPR000436; Sushi_SCR_CCP.
CC InterPro: IPR000130; Zn_MTPetase.
CC Pfam: PF000084; sushi; 4.
CC SMART: SM00032; CCP; 4.
CC SMART: SM00004; NU; 2.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Metal-binding; Zinc; Signal; Glycoprotein;

[illegible]

	Query Match	92.1%	Score 35;	DB 1;	Length 1627;
	Best Local Similarity	83.3%;	Pred. No. 25;		
	Matches	5;	Conservative	1;	Mismatches
				0;	Indels
					Gaps
Oy	1 CVP/LTC	6			
Db	1410 CVP/VTC	1415			

CC	AA2B_MOUSE	STANDARD:	PRT:	332 AA.
AC	Q60614:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Adenosine A2b receptor.			
GN	ADORA2B.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/C; TISSUE-Bone marrow;			
RX	MEDLINE=94209670; PubMed=8157966;			
RA	Marquardt D.L., Walker L.L., Heilmann S.;			
RT	"Cloning of two adenosine receptor subtypes from mouse bone marrow-			
RT	derived mast cells";			
RL	J. Immunol. 152:4508-4515(1994).			
CC	-1- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR			
CC	IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR	EMBL; U05673; AAI19001.1; .	
DR	HSSP; P29274; 1MM.	
DR	GRDB; GCR1024; .	
DR	MGI; MGI:99403; Adora2b.	
DR	InterPro; IPR000276; GPCR_Rhodopsn.	
DR	Pfam; PF00001; 7tm_1; 1..	
DR	PRINTS; PR00237; GPCRHOOPS.	
DR	PROSITE; PS00237; G_PROTEIN_RECCEP_F1_1; 1.	
DR	PROSITE; PS50262; G_PROTEIN_RECCEP_F1_2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;	
KW	Lipoprotein; Palmitate.	
FT	DOMAIN	1
FT	TRANSMEM	9
FT	DOMAIN	32
FT	TRANSMEM	45
FT	DOMAIN	68
FT	TRANSMEM	79
FT	DOMAIN	102
FT	TRANSMEM	122
FT	DOMAIN	145
FT	TRANSMEM	179
FT	DOMAIN	204
FT	TRANSMEM	236
FT	DOMAIN	260
FT	TRANSMEM	268
FT	DOMAIN	292
FT	DISULFID	78
FT	CARBOHYD	153
FT	CARBOHYD	163
FT	LIPID	311
FT	SEQUENCE	332 AA; 36064 MW; C8A3108371ADIICIC CCK64; PALMITATE (POTENTIAL). N-LINKED (GLCNAC...) (POTENTIAL). N-LINKED (GLCNAC...) (POTENTIAL). PALMITATE (POTENTIAL).

		86.8%;	Score 33;	DB 1;	Length 332;
		Best Local Similarity	83.3%;	Pred. No. 14;	
	Matches	5,	Conservative	0;	Mismatches
				1;	Indels
				0;	Gaps
QY	1	CVELTC	6		
		1			
		5			
Db	166	CCPLTC	171		

	RESULT	4
	BM86_BOOMI	
ID	BM86_BOOMI	STANDARD: PRT: 650 AA.
AC	P20736;	
DT	01-FEB-1991 (Rel. 17, Created)	
DT	01-FEB-1991 (Rel. 17, Last sequence update)	
DT	01-NOV-1997 (Rel. 35, Last annotation update)	
DE	Glycoprotein antigen BM86 precursor (Protective antigen).	
OS	Boophilus microplus (Cattle tick).	
OC	Euryarchaeota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;	
OC	Parasitiformes; Ixodida; Ixodidae; Boophilus.	
OX	NCBI_TaxId=6941;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RC	TISSUE=Gut;	
RA	MEDLINE=90099323; PubMed=2690068;	
RA	Rand K.N., Moore T., Srikantha A., Spring K., Tellam R.L.,	
RA	Willadsen P., Cobon G.S.;	
RT	"Cloning and expression of a protective antigen from the cattle tick	
RT	Boophilus microplus.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 86:9657-9661(1989).	
RN	[2]	
RP	PARTIAL SEQUENCE.	
RX	MEDLINE=89309823; PubMed=2745982;	
RX	Willadsen P., Ridling G.A., McKenna R.V., Kemp D.H., Tellam R.L.,	
RT	Nielsen J.N., Lammstein J., Cobon G.S., Gough J.M.;	
RT	"Immunologic control of a parasitic arthropod. Identification of a	

RT protective antigen from *Boophilus microplus*.";
 RL J. Immunol. 143:1346-1351(1989).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Probable).
 CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M29321; AAA30098.1; -
 DR PIR: A34498; A34498.
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF; 2.
 DR SMART: SM00181; EGF; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.
 KW Glycoprotein; Antigen; Signal; EGF-like domain; Repeat; GPI-anchor.
 FT SIGNAL 1 19
 FT CHAIN 20 627
 FT PROPEP 628 650
 FT DOMAIN 20 66
 FT DOMAIN 67 104
 FT DOMAIN 205 247
 FT DOMAIN 251 292
 FT DOMAIN 281 335
 FT DOMAIN 482 530
 FT DOMAIN 531 568
 FT DOMAIN 601 627
 FT DISULFID 24 37
 FT DISULFID 32 49
 FT DISULFID 51 65
 FT DISULFID 71 81
 FT DISULFID 76 91
 FT DISULFID 93 103
 FT DISULFID 103 103
 FT DISULFID 209 222
 FT DISULFID 218 231
 FT DISULFID 233 246
 FT DISULFID 255 269
 FT DISULFID 263 278
 FT DISULFID 280 291
 FT DISULFID 295 307
 FT DISULFID 300 316
 FT DISULFID 318 334
 FT DISULFID 486 500
 FT DISULFID 492 516
 FT DISULFID 518 529
 FT DISULFID 535 550
 FT DISULFID 543 559
 FT DISULFID 561 567
 FT LIPID 627 627
 FT CARBOHYD 141 141
 FT CARBOHYD 182 182
 FT CARBOHYD 348 348
 FT CARBOHYD 382 382
 FT VARIANT 235 237
 FT VARIANT 507 507
 FT SEQUENCE 650 AA; 71721 MW; 5DEB71E354D8312A CRC64;
 Query Match 86.8%; Score 33; DB 1; Length 650;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5
 ID LM3_HUMAN STANDARD; PRT; 1587 AA.
 AC Q9Y6N6;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
 GN LAMC3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Placenta;
 RX MEDLINE-99242614; PubMed-10225960;
 RA Koch M., Olson P.F., Albus A., Jin W., Hunter D.D., Brunken W.J.,
 RA Burgess R.E., Champlaud M.F.;
 RT "Characterization and expression of the laminin gamma3 chain: a novel,
 RT non-basement membrane-associated, laminin chain.";
 RL J. Cell Biol. 145:605-618(1999).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and
 CC the reproductive tracts.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF041835; AM36991.1; -
 DR HSSP: P02468; ITLE.
 DR MIM: 604349; -
 DR InterPro: IPR000561; EGF-like
 DR InterPro: IPR000034; Laminin_B.
 DR InterPro: IPR002049; Laminin_EGF.
 DR InterPro: IPR001886; LAMNT.
 DR Pfam: PF00052; Laminin_B; 1.
 DR Pfam: PF00053; Laminin_EGF; 10.
 DR Pfam: PF00055; Laminin_Nterm; 1.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR ProDom: PD002082; LamNT; 1.
 DR ProDom: PD003031; Laminin_B; 1.
 DR ProDom: PD003031; Laminin_EGF; 10.
 DR SMART: SM00180; EGF_Lam; 10.
 DR SMART: SM00281; EGF_Like; 1.
 DR SMART: SM00281; Lamb; 1.
 DR SMART: SM00136; LAMNT; 1.
 DR PROSITE: PS00022; EGF_1; 7.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01248; LAMININ_Type_EGF; 10.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.

```

FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1587 LAMININ GAMMA-3 CHAIN.
FT DOMAIN 20 270 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 271 326 LAMININ EGF-LIKE 1.
FT DOMAIN 327 382 LAMININ EGF-LIKE 2.
FT DOMAIN 383 429 LAMININ EGF-LIKE 3.
FT DOMAIN 430 479 LAMININ EGF-LIKE 4.
FT DOMAIN 480 489 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 490 672 LAMININ DOMAIN IV.
FT DOMAIN 673 706 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 707 754 LAMININ EGF-LIKE 6.
FT DOMAIN 755 809 LAMININ EGF-LIKE 7.
FT DOMAIN 810 865 LAMININ EGF-LIKE 8.
FT DOMAIN 866 916 LAMININ EGF-LIKE 9.
FT DOMAIN 917 964 LAMININ EGF-LIKE 10.
FT DOMAIN 965 1013 LAMININ EGF-LIKE 11.
FT DOMAIN 1014 1587 DOMAIN II AND I.
FT DOMAIN 1071 1141 COILED COIL (POTENTIAL).
FT DOMAIN 1200 1229 COILED COIL (POTENTIAL).
FT DOMAIN 1424 1504 COILED COIL (POTENTIAL).
FT DOMAIN 1535 1579 COILED COIL (POTENTIAL).
FT SITE 1059 1061 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1185 1185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1518 1518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1587 AA; 172051 MW; 3CB6E09B5F203319 CRC64;

```

```

Query Match      86.8%; Score 33; DB 1; Length 1587;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 CVP LTC 6
DB 704 CVP LTC 709

```

```

RESULT 6
VANB_PSES9 STANDARD; PRT; 314 AA.
AC P12580;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Vanillate O-demethylase oxidoreductase (EC 1.14.13.-) (Vanillate
  degradation ferredoxin-like protein).
GN VANB.
OS Pseudomonas sp. (Strain ATCC 19151).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89008117; PubMed=3170489;
RA Brunel F., Davison J.;
RT "Cloning and sequencing of Pseudomonas genes encoding vanillate
  demethylase."
RT J. Bacteriol. 170:4924-4930(1988).
CC -1- COFACTOR: FMN (BY SIMILARITY).
CC -1- PATHWAY: VANILLATE DEGRADATION (VANILLATE IS A KEY INTERMEDIATE
  IN THE DEGRADATION OF LIGNIN).
CC -1- SIMILARITY: IN THE C-TERMINAL, BELONGS TO THE 2FEZS PLANT-TYPE
  FERREDOXIN FAMILY.
CC -1- SIMILARITY: BELONGS TO THE PDR/VANB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22077; AAA26020.1; -.
DR PIR; B43652; B43652.
DR HSSP; P33164; 2PIA.
DR InterPro; IPR000564; 2FEZS_ferredoxin.
DR InterPro; IPR001834; Cyl_B5_reductase.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001433; Oxidored_FAD.
DR InterPro; IPR000951; Phdiox_reductase.
DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00175; NAD_binding; 1.
DR PRINTS; PR00409; PHDIOXRDTASE.
DR PROSITE; PS00197; 2FEZS_FERREDOXIN; 1.
KW Aromatic hydrocarbons catabolism; Lignin degradation; Oxidoreductase;
  Flavoprotein; FMN; NAD; Iron-sulfur; Electron transport.
FT NP_BIND 1 95 FMN.
FT METAL 105 217 NAD.
FT METAL 263 263 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 268 268 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 271 271 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 302 302 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 314 AA; 33706 MW; FC521516A6CEB72 CRC64;

```

```

Query Match      84.2%; Score 32; DB 1; Length 314;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 CVP LTC 6
DB 258 CVP LTC 263

```

```

RESULT 7
APOH_BOVIN STANDARD; PRT; 345 AA.
AC P17690; Q28052;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-2-glycoprotein I precursor (Apolipoprotein H) (Apo-H) (B2GPI)
  (beta(2)GPI).
GN APOH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX Gao B., Virmani M., Romm E., Lazar-Wesley E., Sakaguchi K.,
  Appella E., Kunos G., Jakacs L.;
  Submitted (DEC-1992) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 4-345 FROM N.A., PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=92232647; PubMed=1567819;
RA Bendisen E., Halkier T., Magnusson S., Sottrup-Jensen L.,
  Kristensen T.;
RT "Complete primary structure of bovine beta 2-glycoprotein I:
  localization of the disulfide bridges."
RT Biochemistry 31:3611-3617(1992).
RN [3]
RP SEQUENCE OF 20-345, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX TISSUE=Plasma;
RX MEDLINE=92089075; PubMed=1751487;
RA Kato H., Enjyoji K.-I.;

```

RT "Amino acid sequence and location of the disulfide bonds in bovine
beta 2 glycoprotein I: the presence of five Sush1 domains.";
RL Biochemistry 30:11687-11694(1991).
RN [4]
RP SEQUENCE OF 20-41.
RX MEDLINE=90226328; PubMed=2327984;
RA L.A. O., Blacher R., Esch F., Congote L.F.;
RT "Isolation from fetal bovine serum of an apolipoprotein-H-like
protein which inhibits thymidine incorporation in fetal calf
erythroid cells";
RL Biochem. J. 267:261-264(1990).
CC -1- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: CONTAINS 4 SUSH1 (SCR) DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L07303; AAA30382.1; -;
DR EMBL: X60065; CAA42669.1; -;
DR PIR: S23597; NBBO.
DR PIR: S09032; S09032.
DR HSSP: P02749; 1C12.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1; 4.
DR SMART: SM00032; CCP; 4.
KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sush1; Signal.
FT SIGNAL 1 19
FT CHAIN 1 19
FT DOMAIN 20 345 BETA-2-GLYCOPROTEIN I.
FT DOMAIN 22 80 SUSH1 1.
FT DOMAIN 83 138 SUSH1 2.
FT DOMAIN 141 201 SUSH1 3.
FT DOMAIN 204 261 SUSH1 4.
FT DOMAIN 263 345 SUSH1-LIKE.
FT DISULFID 23 66
FT DISULFID 51 79
FT DISULFID 84 124
FT DISULFID 110 137
FT DISULFID 142 188
FT DISULFID 174 200
FT DISULFID 205 248
FT DISULFID 234 260
FT DISULFID 264 315
FT DISULFID 300 325
FT DISULFID 307 345
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .)
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .)
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .)
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .)
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .)
FT CONFLICT 101 101 E -> G (IN REF. 1).
FT CONFLICT 108 108 E -> S (IN REF. 1).
FT CONFLICT 177 177 H -> R (IN REF. 1).
FT CONFLICT 191 191 H -> N (IN REF. 3).
FT CONFLICT 194 194 W -> C (IN REF. 1).
FT CONFLICT 259 259 S -> N (IN REF. 1).
FT CONFLICT 302 302 H -> N (IN REF. 1).
FT CONFLICT 305 305 K -> R (IN REF. 1).
FT CONFLICT 329 329 H -> R (IN REF. 1).
SQ SEQUENCE 345 AA; 38252 MW; E117DAB609461C33 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 345;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVP/LTC 6
Db 137 CAP/LTC 142

RESULT 8
DCAM_TRYBB STANDARD; PRT; 370 AA.
ID DCAM_TRYBB
AC P50244;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdoMetDC)
DE (Samc) [contains: S-adenosylmethionine decarboxylase alpha chain; S-
DE adenosylmethionine decarboxylase beta chain].
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FAIRO 164;
RA Scott J.R., Ullman B.;
RT "Molecular cloning and functional expression of the S-
RT adenosylmethionine decarboxylase gene of Leishmania donovani and
RT Trypanosoma brucei.";
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine - (5-deoxy-5-
CC adenosyl)(3-aminopropyl) methylsulfonium salt + CO(2).
CC -1- COFACTOR: PYRUVYL GROUP.
CC -1- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
CC AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE
CC BIOSYNTHESIS FROM PUTRESCINE.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC ADOMETDC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20092; AAA61969.1; -;
DR HSSP: P17707; LJEN.
DR InterPro: IPR001985; SAM_decarbox.
DR Pfam: PF01536; SAM_decarbox; 1.
DR ProDom: PD002379; SAM_decarbox; 1.
DR PROSITE: PS01336; ADOMETDC; 1.
KW Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen.
FT CHAIN 1 85
FT CHAIN 86 370
FT SITE 85 86 CHAIN (BY SIMILARITY).
FT SITE 85 86 CLEAVAGE (NONHYDROLYTIC)
FT MOD_RES 86 86 CONVERTED TO A PYRUVYL GROUP
FT ACT_SITE 29 29 (BY SIMILARITY).
FT ACT_SITE 29 29 IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT ACT_SITE 32 32 SIMILARITY).
FT ACT_SITE 32 32 IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT ACT_SITE 100 100 SIMILARITY).
FT ACT_SITE 100 100 IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT ACT_SITE 100 100 SIMILARITY).
SQ SEQUENCE 370 AA; 41692 MW; 06E0BF4BEAD4D42 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 370;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVP/LTC 6
Db 137 CAP/LTC 142

Db 108 CVPLOC 113

RESULT 9

MOBL_XENLA STANDARD; PRT; 398 AA.

AC P38565;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Integumentary mucin B.1 (FIM-B.1) (Fragment).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP MEDLINE-91002513; PubMed-2207068;

RA Probst J.C., Gertzen E.-M., Hoffmann W.;

RT "An integumentary mucin (FIM-B.1) from Xenopus laevis homologous with von Willebrand factor.";

RL Biochemistry 29:6240-6244(1990).

CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.

CC -1- TISSUE SPECIFICITY: EXPRESSED AND STORED EXCLUSIVELY IN MATURE MUCOUS GLANDS OF THE SKIN.

CC -1- PFM: EXTENSIVELY O-GLYCOSYLATED.

CC -1- SIMILARITY: CONTAINS 1 WMFC DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J02910; AAA9711.1; -.

DR PIR: A35281; A35281.

DR HSP: P04355; 4MT2.

DR InterPro: IPR000359; Cys_knot.

DR InterPro: IPR001007; WMFC.

DR pfam: PR00007; Cys_knot; 1.

DR SMART: SM00041; CT_1.

DR SMART: SM00214; VWC; 1.

DR PROSITE: PS01208; WMFC; 1.

DR PROSITE: PS01185; CTCK_1; 1.

DR PROSITE: PS01225; CTCK_2; 1.

KW Repeat: Amphibian skin; Glycoprotein.

FT NON_TER 1 1

FT DOMAIN 1 93

FT REPEAT <1 7 8 X 11 AA TANDEM REPEATS OF G-E-S-T-P-

FT REPEAT 8 18 A-P-S-E-T-T.

FT REPEAT 19 29 1 (INCOMPLETE).

FT REPEAT 30 40

FT REPEAT 41 51

FT REPEAT 52 62

FT REPEAT 66 76

FT REPEAT 83 93

FT REPEAT 170 234

FT DOMAIN 170 234

FT DOMAIN 301 356

FT DISULFID 301 356

FT DISULFID 323 370

FT DISULFID 332 386

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

FT DISULFID 336 388

Query Match 84.2%; Score 32; DB 1; Length 398;
Best Local Similarity 83.3%; Pred. NO. 26;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;QY 1 CVPLOC 6
1 11111
Db 233 CVPLOC 238

RESULT 10

CHRD_BRARE

ID CHRD_BRARE STANDARD; PRT; 940 AA.

AC 057472; Q9DEB8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Chordin precursor (Chordin protein).

GN CHD.

OS Brachydanio rerio (zebrafish) (zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RP MEDLINE-98104254; PubMed-9441687;

RA Muller-Bertoglio V.E., Fisher S., Sanchez A., Mullins M.C.,

RT "Differential regulation of chordin expression domains in mutant zebrafish.";

RL Dev. Biol. 192:537-550(1997).

RN [2]

RP SEQUENCE OF 1-42 FROM N.A.

RA Fujii R., Hibi M., Hirano T., Shimizu T.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Dorsalizing factor. Key developmental protein that dorsalizes early vertebrate embryonic tissues by binding to

CC ventralizing TGF-beta family bone morphogenetic proteins (Bmps) and sequestering them in latent complexes (by similarity).

CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).

CC -1- TISSUE SPECIFICITY: During gastrulation, levels are highest in the organizer region. Also present in the developing brain and in

CC paraxial mesoderm and ectoderm.

CC -1- DEVELOPMENTAL STAGE: First detected shortly after the midblastula transition. Levels increase during gastrulation, persist through

CC early somitogenesis, but then decrease and are gone by 24

CC hours.

CC -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.

CC -1- SIMILARITY: CONTAINS 4 WMFC DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF034606; AAB93485.1; -.

DR EMBL: AB043968; BAB18642.1; -.

DR ZFIN: ZDB-GENE-990415-33; chd.

DR InterPro: IPR001007; WMFC.

DR pfam: PR00093; VWC; 4.

DR SMART: SM00214; VWC; 4.

DR PROSITE: PS01208; WMFC; 3.

KW Developmental protein; Repeat; Glycoprotein; Signal.

FT SIGNAL 1 19

FT CHAIN 20 940

FT DOMAIN 42 118

FT DOMAIN 689 748

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 767 836

FT DOMAIN 855 919 VMFC 4.
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 940 AA; 104999 MW; B855CAF84F4623AC CRC64;

 Query Match 84.2%; Score 32; DB 1; Length 940;
 Best Local Similarity 83.3%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 Oy 1 CVP LTC 6
 Db 815 CVP LTC 820

 RESULT 11
 LMG3_MOUSE STANDARD; PRT; 1581 AA.
 AC Q9R0B6; Q9R0T6; 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
 GN LAMC3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 ON NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Aldus A.M., Burgess B., Champilaud M.-F., Koch M., Olson P.;
 RT "Mouse Laminin 12 gamma 3 chain."
 RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-1526 FROM N.A.
 RX MEDLINE=99253969; PubMed=10318827;
 RA Iivanainen A., Morita T., Tryggvason K.;
 RT "Molecular cloning and tissue-specific expression of a novel murine
 RL Laminin gamma3 chain."
 RL J. Biol. Chem. 274:14107-14111(1999).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Strongly expressed in capillaries and
 CC arterioles of kidney as well as in interstitial Leydig cells of
 CC testis.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 11 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF083372; AAF08983.1; -;
 DR EMBL; AF079520; AAD29851.1; -;
 DR HSSP; P02468; IKLO.
 DR MGD; MGI:1344394; Lamc3.
 DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001886; LamNT.
 DR Pfam; PF00052; Laminin_B; 1.
 DR Pfam; PF00053; Laminin_EGF; 10.
 DR Pfam; PF00055; Laminin_Nterm; 1.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR ProDom; PD002082; LamNT; 1.
 DR ProDom; PD003031; Laminin_B; 1.
 DR SMART; SM00180; EGF_Lam; 10.
 DR SMART; SM00281; Lamb; 1.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_8.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_Type_EGF; 11.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 1581
 FT DOMAIN 29 279 LAMININ GAMMA-3 CHAIN.
 FT DOMAIN 29 335 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 336 391 LAMININ EGF-LIKE 1.
 FT DOMAIN 392 391 LAMININ EGF-LIKE 2.
 FT DOMAIN 392 438 LAMININ EGF-LIKE 3.
 FT DOMAIN 439 488 LAMININ EGF-LIKE 4.
 FT DOMAIN 489 498 LAMININ EGF-LIKE 5 (N-TERMINAL).
 FT DOMAIN 499 684 LAMININ DOMAIN IV.
 FT DOMAIN 685 718 LAMININ EGF-LIKE 5 (C-TERMINAL).
 FT DOMAIN 719 766 LAMININ EGF-LIKE 6.
 FT DOMAIN 767 821 LAMININ EGF-LIKE 7.
 FT DOMAIN 822 877 LAMININ EGF-LIKE 8.
 FT DOMAIN 878 927 LAMININ EGF-LIKE 9.
 FT DOMAIN 928 975 LAMININ EGF-LIKE 10.
 FT DOMAIN 976 1024 LAMININ EGF-LIKE 11.
 FT DOMAIN 1025 1581 DOMAIN II AND I.
 FT DOMAIN 1029 1046 COILED COIL (POTENTIAL).
 FT DOMAIN 1112 1133 COILED COIL (POTENTIAL).
 FT DOMAIN 1208 1231 COILED COIL (POTENTIAL).
 FT DOMAIN 1438 1468 COILED COIL (POTENTIAL).
 FT DOMAIN 1510 1575 COILED COIL (POTENTIAL).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 849 849 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1162 1162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1514 1514 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 9 9
 FT CONFLICT 190 190 L -> F (IN REF. 2).
 FT CONFLICT 195 195 R -> T (IN REF. 2).
 FT CONFLICT 221 221 R -> K (IN REF. 2).
 FT CONFLICT 221 221 G -> S (IN REF. 2).
 FT CONFLICT 394 394 C -> Y (IN REF. 2).
 FT CONFLICT 471 471 C -> Y (IN REF. 2).
 FT CONFLICT 1150 1150 L -> IDEPQFSLLK (IN REF. 2).
 FT CONFLICT 1387 1387 H -> Q (IN REF. 2).
 FT CONFLICT 1438 1439 AS -> TI (IN REF. 2).
 FT CONFLICT 1479 1479 V -> I (IN REF. 2).
 SO SEQUENCE 1581 AA; 172316 MW; 51DFADLF95E6AD81 CRC64;

 Query Match 84.2%; Score 32; DB 1; Length 1581;
 Best Local Similarity 66.7%; Pred. No. 95;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 Oy 1 CVP LTC 6
 Db 716 CVP LTC 721

 RESULT 12
 LRP2_HUMAN


```

FT DOMAIN 1655 1695 LDL-RECEPTOR CLASS B 14.
FT DOMAIN 1700 1741 EGF-LIKE 7.
FT DOMAIN 1790 1831 LDL-RECEPTOR CLASS B 15.
FT DOMAIN 1833 1881 LDL-RECEPTOR CLASS B 16.
FT DOMAIN 1883 1929 LDL-RECEPTOR CLASS B 17.
FT DOMAIN 1931 1971 LDL-RECEPTOR CLASS B 18.
FT DOMAIN 1972 2012 LDL-RECEPTOR CLASS B 19.
FT DOMAIN 2018 2059 EGF-LIKE 8.
FT DOMAIN 2107 2155 LDL-RECEPTOR CLASS B 20.
FT DOMAIN 2157 2200 LDL-RECEPTOR CLASS B 21.
FT DOMAIN 2202 2244 LDL-RECEPTOR CLASS B 22.
FT DOMAIN 2246 2288 LDL-RECEPTOR CLASS B 23.
FT DOMAIN 2290 2331 LDL-RECEPTOR CLASS B 24.
FT DOMAIN 2342 2383 EGF-LIKE 9.
FT DOMAIN 2431 2476 LDL-RECEPTOR CLASS B 25.
FT DOMAIN 2478 2517 LDL-RECEPTOR CLASS B 26.
FT DOMAIN 2519 2561 LDL-RECEPTOR CLASS B 27.
FT DOMAIN 2563 2603 LDL-RECEPTOR CLASS B 28.
FT DOMAIN 2604 2646 LDL-RECEPTOR CLASS B 29.
FT DOMAIN 2651 2693 EGF-LIKE 10.
FT DOMAIN 2698 2738 LDL-RECEPTOR CLASS A 16.
FT DOMAIN 2739 2777 LDL-RECEPTOR CLASS A 17.
FT DOMAIN 2778 2819 LDL-RECEPTOR CLASS A 18.
FT DOMAIN 2820 2861 LDL-RECEPTOR CLASS A 19.
FT DOMAIN 2862 2901 LDL-RECEPTOR CLASS A 20.
FT DOMAIN 2904 2945 LDL-RECEPTOR CLASS A 21.
FT DOMAIN 2946 2990 LDL-RECEPTOR CLASS A 22.
FT DOMAIN 2991 3029 LDL-RECEPTOR CLASS A 23.
FT DOMAIN 3030 3070 LDL-RECEPTOR CLASS A 24.
FT DOMAIN 3073 3110 LDL-RECEPTOR CLASS A 25.
FT DOMAIN 3111 3151 EGF-LIKE 11.
FT DOMAIN 3152 3192 EGF-LIKE 12.
FT DOMAIN 3239 3280 LDL-RECEPTOR CLASS B 30.
FT DOMAIN 3282 3331 LDL-RECEPTOR CLASS B 31.
FT DOMAIN 3333 3375 LDL-RECEPTOR CLASS B 32.
FT DOMAIN 3377 3418 LDL-RECEPTOR CLASS B 33.
FT DOMAIN 3419 3459 LDL-RECEPTOR CLASS B 34.
FT DOMAIN 3465 3509 EGF-LIKE 13.
FT DOMAIN 3510 3550 LDL-RECEPTOR CLASS A 26.
FT DOMAIN 3551 3591 LDL-RECEPTOR CLASS A 27.
FT DOMAIN 3592 3632 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3633 3673 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3676 3716 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3717 3756 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3757 3795 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3834 3880 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3881 3922 LDL-RECEPTOR CLASS A 34.
FT DOMAIN 3926 3964 LDL-RECEPTOR CLASS A 35.
FT DOMAIN 3966 4006 EGF-LIKE 14.
FT DOMAIN 4007 4048 EGF-LIKE 15.
FT DOMAIN 4048 4195 LDL-RECEPTOR CLASS B 35.
FT DOMAIN 4157 4239 LDL-RECEPTOR CLASS B 36.
FT DOMAIN 4242 LDL-RECEPTOR CLASS B 37.

```

```

Query Match      84.2%   Score 32; DB 1; Length 4655;
Best Local Similarity 83.3%   Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 CYP11C 6
Db 1266 CYP11C 1271

```

```

RESULT 13
LRP2-RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
   (Glycoprotein 330).

```

```

GN LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Plethronaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/'megalin,' a
RT distinctive member of the low density lipoprotein receptor gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RN [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gilemann J., Christensen E.I.;
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of
RT polybasic drugs.";
RL J. Clin. Invest. 96:1404-1413(1995).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
RA Andres G., McCluskey R.T.;
RT "Organ distribution in rats of two members of the low-density
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the
RT receptor-associated protein (RAP).";
RL J. Histochem. Cytochem. 42:531-542(1994).
CC -1- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
CC PLASMINOGEN ACTIVATOR-PLASMINOGEN ACTIVATOR INHIBITOR TYPE I
CC COMPLEX, APOLOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,
CC LACTOFERRIN CLUSTERIN AND CALCIUM.
CC -1- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
CC APORINTIN, AMINOGLYCOSIDES AND POLYMYXIN B.
CC -1- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
CC ASSOCIATED PROTEIN (RAP).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN
CC CLATHRIN-COATED PITS: A SOLUBLE FORM IS POSSIBLY DERIVED BY
CC CLEAVAGE AT THE CELL SURFACE.
CC -1- TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND
CC PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.
CC -1- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L34049; AAA51369.1; -
DR HSSP: Q07954; ICR8.
DR GICONSULEDB: P98158; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000033; LDL_receptor_rep.
DR Pfam: PF00057; ldl_recept_a; 36.
DR Pfam: PF00058; ldl_recept_b; 33.
DR PRINTS: PR00261; LDLRECEPTOR.
DR SMART: SM00179; EGF_CA_3.
DR SMART: SM00001; EGF_LIKE_3.
DR SMART: SM00019; LDLa; 36.
DR SMART: SM00135; LY; 35.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; 1.

```

DR	PROSITE: PS01186; EGF_CA; 8.	
DR	PROSITE: PS01187; EGF_CA; 31.	
DR	PROSITE: PS01209; LDLRA_1; 31.	
DR	PROSITE: PS01068; LDLRA_2; 36.	
KW	Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;	
KW	glycosylator; EGF-like domain; SH3-binding; Signal.	
FT	SIGNAL	1 25
FT	CHAIN	26 4660
FT		
FT	DOMAIN	26 4425
FT	TRANSMEM	4426 4446
FT	DOMAIN	4447 4660
FT	DOMAIN	26 65
FT	DOMAIN	65 105
FT	DOMAIN	106 144
FT	DOMAIN	145 181
FT	DOMAIN	181 219
FT	DOMAIN	220 258
FT	DOMAIN	263 307
FT	DOMAIN	308 346
FT	DOMAIN	347 385
FT	DOMAIN	435 476
FT	DOMAIN	478 519
FT	DOMAIN	521 566
FT	DOMAIN	568 611
FT	DOMAIN	612 652
FT	DOMAIN	658 704
FT	DOMAIN	752 793
FT	DOMAIN	795 835
FT	DOMAIN	837 879
FT	DOMAIN	881 923
FT	DOMAIN	969 1013
FT	DOMAIN	1023 1061
FT	DOMAIN	1064 1103
FT	DOMAIN	1108 1146
FT	DOMAIN	1148 1186
FT	DOMAIN	1186 1225
FT	DOMAIN	1229 1269
FT	DOMAIN	1270 1308
FT	DOMAIN	1311 1351
FT	DOMAIN	1350 1390
FT	DOMAIN	1391 1430
FT	DOMAIN	1479 1520
FT	DOMAIN	1522 1563
FT	DOMAIN	1567 1609
FT	DOMAIN	1611 1654
FT	DOMAIN	1656 1696
FT	DOMAIN	1701 1742
FT	DOMAIN	1791 1832
FT	DOMAIN	1834 1882
FT	DOMAIN	1884 1930
FT	DOMAIN	1932 1972
FT	DOMAIN	1973 2013
FT	DOMAIN	2019 2060
FT	DOMAIN	2108 2156
FT	DOMAIN	2158 2201
FT	DOMAIN	2203 2245
FT	DOMAIN	2247 2289
FT	DOMAIN	2291 2332
FT	DOMAIN	2343 2384
FT	DOMAIN	2432 2477
FT	DOMAIN	2479 2518
FT	DOMAIN	2520 2562
FT	DOMAIN	2564 2604
FT	DOMAIN	2605 2647
FT	DOMAIN	2652 2694
FT	DOMAIN	2699 2739
FT	DOMAIN	2740 2778
FT	DOMAIN	2779 2820
FT	DOMAIN	2821 2862
FT	DOMAIN	2863 2903
FT	DOMAIN	2906 2947
FT	DOMAIN	2948 2992
DR	PROSITE: PS01186; EGF_CA; 8.	
DR	PROSITE: PS01187; EGF_CA; 31.	
DR	PROSITE: PS01209; LDLRA_1; 31.	
DR	PROSITE: PS01068; LDLRA_2; 36.	
KW	Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;	
KW	glycosylator; EGF-like domain; SH3-binding; Signal.	
FT	SIGNAL	1 25
FT	CHAIN	26 4660
FT		
FT	DOMAIN	26 4425
FT	TRANSMEM	4426 4446
FT	DOMAIN	4447 4660
FT	DOMAIN	26 65
FT	DOMAIN	65 105
FT	DOMAIN	106 144
FT	DOMAIN	145 181
FT	DOMAIN	181 219
FT	DOMAIN	220 258
FT	DOMAIN	263 307
FT	DOMAIN	308 346
FT	DOMAIN	347 385
FT	DOMAIN	435 476
FT	DOMAIN	478 519
FT	DOMAIN	521 566
FT	DOMAIN	568 611
FT	DOMAIN	612 652
FT	DOMAIN	658 704
FT	DOMAIN	752 793
FT	DOMAIN	795 835
FT	DOMAIN	837 879
FT	DOMAIN	881 923
FT	DOMAIN	969 1013
FT	DOMAIN	1023 1061
FT	DOMAIN	1064 1103
FT	DOMAIN	1108 1146
FT	DOMAIN	1148 1186
FT	DOMAIN	1186 1225
FT	DOMAIN	1229 1269
FT	DOMAIN	1270 1308
FT	DOMAIN	1311 1351
FT	DOMAIN	1350 1390
FT	DOMAIN	1391 1430
FT	DOMAIN	1479 1520
FT	DOMAIN	1522 1563
FT	DOMAIN	1567 1609
FT	DOMAIN	1611 1654
FT	DOMAIN	1656 1696
FT	DOMAIN	1701 1742
FT	DOMAIN	1791 1832
FT	DOMAIN	1834 1882
FT	DOMAIN	1884 1930
FT	DOMAIN	1932 1972
FT	DOMAIN	1973 2013
FT	DOMAIN	2019 2060
FT	DOMAIN	2108 2156
FT	DOMAIN	2158 2201
FT	DOMAIN	2203 2245
FT	DOMAIN	2247 2289
FT	DOMAIN	2291 2332
FT	DOMAIN	2343 2384
FT	DOMAIN	2432 2477
FT	DOMAIN	2479 2518
FT	DOMAIN	2520 2562
FT	DOMAIN	2564 2604
FT	DOMAIN	2605 2647
FT</		

FT	DOMAIN	2293	3031	LDL-RECEPTOR CLASS A 23.
FT	DOMAIN	3032	3072	LDL-RECEPTOR CLASS A 24.
FT	DOMAIN	3075	3112	LDL-RECEPTOR CLASS A 25.
FT	DOMAIN	3113	3153	EGF-LIKE 11.
FT	DOMAIN	3154	3194	EGF-LIKE 12. CALCULUM-BINDING (POTENTIAL).
FT	DOMAIN	3241	3282	LDL-RECEPTOR CLASS B 30.
FT	DOMAIN	3284	3333	LDL-RECEPTOR CLASS B 31.
FT	DOMAIN	3335	3377	LDL-RECEPTOR CLASS B 32.
FT	DOMAIN	3379	3420	LDL-RECEPTOR CLASS B 33.
FT	DOMAIN	3421	3461	LDL-RECEPTOR CLASS B 34.
FT	DOMAIN	3467	3511	EGF-LIKE 13.
FT	DOMAIN	3512	3552	LDL-RECEPTOR CLASS A 26.
FT	DOMAIN	3553	3593	LDL-RECEPTOR CLASS A 27.
FT	DOMAIN	3594	3634	LDL-RECEPTOR CLASS A 28.
FT	DOMAIN	3635	3675	LDL-RECEPTOR CLASS A 29.
FT	DOMAIN	3678	3718	LDL-RECEPTOR CLASS A 30.
FT	DOMAIN	3719	3758	LDL-RECEPTOR CLASS A 31.
FT	DOMAIN	3759	3797	LDL-RECEPTOR CLASS A 32.
FT	DOMAIN	3798	3836	LDL-RECEPTOR CLASS A 33.
FT	DOMAIN	3842	3882	LDL-RECEPTOR CLASS A 34.
FT	DOMAIN	3883	3924	LDL-RECEPTOR CLASS A 35.
FT	DOMAIN	3928	3966	LDL-RECEPTOR CLASS A 36.
FT	DOMAIN	3968	4008	EGF-LIKE 14.
FT	DOMAIN	4009	4050	EGF-LIKE 15. CALCULUM-BINDING (POTENTIAL).
FT	DOMAIN	4156	4197	LDL-RECEPTOR CLASS B 35.
FT	DOMAIN	4199	4241	LDL-RECEPTOR CLASS B 36.
FT	DOMAIN	4244	4284	LDL-RECEPTOR CLASS B 37.
FT	DOMAIN	4332	4370	EGF-LIKE 16.
FT	DOMAIN	4379	4413	EGF-LIKE 17.
FT	SITE	4454	4460	SH3-BINDING (POTENTIAL).
FT	SITE	4457	4463	SH3-BINDING (POTENTIAL).
FT	SITE	4606	4609	SH2-BINDING (POTENTIAL).
FT	SITE	4619	4625	SH3-BINDING (POTENTIAL).
FT	SITE	4624	4630	SH3-BINDING (POTENTIAL).
FT	SITE	1743	1745	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	4422	4527	ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	SITE	4601	4606	ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	DISULFID	28	40	BY SIMILARITY.
FT	DISULFID	35	53	BY SIMILARITY.
FT	DISULFID	47	62	BY SIMILARITY.
FT	DISULFID	67	80	BY SIMILARITY.
FT	DISULFID	74	93	BY SIMILARITY.
FT	DISULFID	87	103	BY SIMILARITY.
FT	DISULFID	108	120	BY SIMILARITY.
FT	DISULFID	115	133	BY SIMILARITY.
FT	DISULFID	127	142	BY SIMILARITY.
FT	DISULFID	147	157	BY SIMILARITY.
Query Match				
Best Local Similarity		84.2%;	Score 32; DB 1; Length 4660;	
Matches		5; Conservative	0; Mismatches	1; Indels
				0; Gaps
QY	1 CVPKTC 6			
Qb	1267 CVPKTC 1272			
RESULT 14				
ID	BD05_BOVIN	STANDARD;	PRT;	64 AA.
AC	P46163; O97533;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-defensin 5 precursor (BNDB-5) (BNDB-5).			
GN	DEFB5 OR BNBD5.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_Taxid=9913;			
NP	[1]			
SEQUENCE FROM N.A.				

RA Kurts B., Pitra C., Schwerin M., Seyfert H.M.;
 RT "beta defensin-encoding genes are selected for divergent sequences of
 the mature anti-bacterial peptide.";
 RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE-ALVEOLAR macrophage;
 RX MEDLINE=98114406; PubMed=9453661;
 RA Ryan L.K., Rhodes J., Bhat M., Diamond G.;
 RT "Expression of beta-defensin genes in bovine alveolar macrophages.";
 RL Infect. Immun. 66:878-881(1998).
 RN [3]
 RP SEQUENCE OF 23-64, FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=HEREFORD; TISSUE=Neutrophils;
 RX MEDLINE=93203264; PubMed=8454635;
 RA Seisted M.E., Tang X.-Q., Morris W.L., McGuire P.A., Novotny M.J.,
 RT Smith W., Henschen A.H., Cullor J.S.;
 RT "Purification, primary structures, and antibacterial activities of
 beta-defensins, a new family of antimicrobial peptides from bovine
 neutrophils.";
 RL J. Biol. Chem. 268:6641-6648(1993).
 RN [4]
 RP REVISIONS TO C-TERMINUS.
 RA Seisted M.E.;
 RL Submitted (May-1996) to the SWISS-PROT data bank.
 CC -1- FUNCTION: HAS BACTERICIDAL ACTIVITY. ACTIVE AGAINST E.COLI ML35
 CC BUT NOT AGAINST S.AUREUS 502A.
 CC -1- TISSUE SPECIFICITY: NEUTROPHILIC GRANULES. ALVEOLAR MACROPHAGES.
 CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ278799; CAC15400.1; -;
 DR HSSP; AF014108; AAD01523.1; -;
 DR HSSP; P46170; 1BNB.
 DR InterPro: IPR001271; Defensin.
 DR InterPro: IPR001855; Defensin_beta.
 DR Pfam: PF00711; Defensin_beta; 1.
 DR SMART; SMO0048; DEFNS; 1.
 DR AntiBiotic; Signal.
 FT SIGNAL 1 22
 FT CHAIN 1 64 BETA-DEFENSIN 5
 FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 31 60 BY SIMILARITY.
 FT DISULFID 38 53 BY SIMILARITY.
 FT DISULFID 43 61 BY SIMILARITY.
 FT CONFLICT 54 54 F -> S (IN REF. 2).
 SO SEQUENCE 64 AA; 7228 MW; 125A5278709131FC CRC64;

Query Match 81.6%; Score 31; DB 1; Length 64;
 Best Local Similarity 50.0%; Pred. No. 7.7;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLFC 6
 DB 38 CIPISC 43

RESULT 15
 OB_SMICR
 ID OB_SMICR STANDARD; PRT; 167 AA.
 AC Q9XSW9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leptin precursor (Obesity factor).

GN LEP OR OB.
 OS Smnithopsis crassicaudata (Fat-tailed dunnart).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Smnithopsis.
 OX NCBI_TaxID=9301;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hope P.J., Lok S., Webb G.C., Hope R.M., Turnbull H., Wilttert G.A.;
 RT "Cloning of leptin cDNA and assignment to the long arm of chromosome 5
 in the marsupial Smnithopsis crassicaudata.";
 RL Submitted (JUN-1993) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS TO
 CC REGULATE THE SIZE OF THE BODY FAT DEPOSIT. AN INCREASE IN THE LEVEL
 CC OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT FOOD
 CC INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A HOMEOSTATIC
 CC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF159713; AAD44337.1; -;
 DR HSSP; P41159; 1AX8.
 DR InterPro: IPR000065; Leptin.
 DR Pfam: PF02024; Leptin; 1.
 DR PRINTS; PR00495; LEPTIN.
 KW Obesity; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 167 POTENTIAL.
 FT DISULFID 117 167 LEPTIN
 FT DISULFID 117 167 BY SIMILARITY.
 SO SEQUENCE 167 AA; 18891 MW; 4825556944B69446 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 167;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPLFC 6
 DB 3 CVPLFC 8

Search completed: June 17, 2002, 16:24:31
 Job time: 1560 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:42 ; Search time 172.85 Seconds
(without alignments)
6.005 Million cell updates/sec

Title: US-09-761-636A-12
Perfect score: 38
Sequence: 1 CVP LTC 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	97.4	407	5	018280 caenorhabdi
2	37	97.4	447	5	0963E7 caenorhabdi
3	36	94.7	533	4	09H4W4 homa sapien
4	36	94.7	3389	4	096Q09 homa sapien
5	36	94.7	3508	4	096RM4 homa sapien
6	36	94.7	3564	11	092313 mus musculu
7	35	92.1	426	13	09DGM2 brachydantio
8	35	92.1	468	11	09JMS7 mus musculu
9	35	92.1	946	4	096Q03 homo sapien
10	35	92.1	1367	11	09ES06 mus musculu
11	34	89.5	367	5	09XU36 caenorhabdi
12	33	86.8	66	13	09PTM2 coturnix co
13	33	86.8	66	16	09A1P4 streptococ
14	33	86.8	68	13	09PTM3 gallus gall
15	33	86.8	93	15	078161 human immun
16	33	86.8	168	13	090YB8 brachydantio

17	33	86.8	246	5	095YD9	095YD9 caenorhabdi
18	33	86.8	277	12	09PYV4	09PYV4 xestia c-ni
19	33	86.8	395	10	09LID1	09LID1 arabidopsis
20	33	86.8	415	13	0915A8	0915A8 xenopus lae
21	33	86.8	429	13	091650	091650 xenopus lae
22	33	86.8	637	10	049077	049077 cucumis mel
23	33	86.8	637	10	09SSY5	09SSY5 cucumis sat
24	33	86.8	650	5	09YOV1	09YOV1 boophilus m
25	33	86.8	1007	13	09OZM3	09OZM3 gallus gall
26	33	86.8	1230	5	024284	024284 drosophila
27	33	86.8	1236	5	09YXN3	09YXN3 drosophila
28	32	84.2	109	11	09D7S3	09D7S3 mus musculu
29	32	84.2	137	5	09XWR9	09XWR9 caenorhabdi
30	32	84.2	156	16	09PP24	09PP24 campylobact
31	32	84.2	252	16	09PCX8	09PCX8 xyella fas
32	32	84.2	301	3	094518	094518 schizosacch
33	32	84.2	345	6	09SLB0	09SLB0 pan troglod
34	32	84.2	382	2	09ETP3	09ETP3 corynebacte
35	32	84.2	425	3	094669	094669 schizosacch
36	32	84.2	447	10	004114	004114 perilla fru
37	32	84.2	467	11	09R0X2	09R0X2 mus musculu
38	32	84.2	544	3	09HGZ7	09HGZ7 schizosacch
39	32	84.2	742	5	09W322	09W322 drosophila
40	32	84.2	843	15	090096	090096 human immun
41	32	84.2	917	13	098931	098931 gallus gall
42	32	84.2	926	5	018504	018504 sepiella offic
43	32	84.2	988	6	097867	097867 sus scrofa
44	32	84.2	1023	5	0967H9	0967H9 strongyloce
45	32	84.2	1034	11	035888	035888 rattus norv

ALIGNMENTS

RESULT 1
ID 018280 PRELIMINARY: PRT: 407 AA.
AC 018280;
DC 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ZC504.5 PROTEIN.
GN ZC504.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99059613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z50029; CAA90343.2; -.
SQ SEQUENCE 407 AA; 46687 MW; 00F390B7952D12B5 CRC64;

Query Match 97.4% Score 37; DB 5; Length 407;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVP LTC 6
Db 262 CVP LTC 267

RESULT 2
Q963E7

ID 0963E7 PRELIMINARY; PRT: 447 AA.
 AC 0963E7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE CHEMORECEPTOR GUR-3.
 GN GUR-3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 RX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Robertson H.M.;
 RT "Gustatory related receptors in nematodes."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF387606; AAK70489.1; -
 KW Receptor.
 SO SEQUENCE 447 AA; 51176 MW; 3AB7B4216F048BBB CRC64;

Query Match 97.4%; Score 37; DB 5; Length 447;
 Best Local Similarity 83.3%; Pred. No. 4;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 302 CIPITC 307

RESULT 3
 ID 094W4 PRELIMINARY; PRT: 533 AA.
 AC 094W4;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DJ947L8.1.8 (NOVEL SUSHI (SCR REPEAT) DOMAIN PROTEIN) (FRAGMENT).
 GN DJ947L8.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis J.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL355178; CAC10283.1; -
 DR HSSP: P10998; 1YVD.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00084; sush1.8.
 DR SMART: SM00032; CCP; 8.
 FT NON_TER 1
 FT NON_TER 533
 SO SEQUENCE 533 AA; 57921 MW; 0554F7E10911F9BF CRC64;

Query Match 94.7%; Score 36; DB 4; Length 533;
 Best Local Similarity 83.3%; Pred. No. 7.6;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 353 CVPITC 358

RESULT 4
 ID 096U9 PRELIMINARY; PRT: 3389 AA.
 AC 096U9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE CUB AND SUSHI MULTIPLE DOMAINS PROTEIN 1 SHORT FORM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21365705; PubMed=11472063;
 RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
 RA Sunwoo J.B., Gollin S.M., Scholnick S.B.;
 RT "Transcript map of the 8p23 putative tumor suppressor region."
 RL Genomics 75:17-25(2001).
 DR EMBL: AY017307; AAG52948.1; -
 SO SEQUENCE 3389 AA; 370293 MW; 53C309FCD3ED76D CRC64;

Query Match 94.7%; Score 36; DB 4; Length 3389;
 Best Local Similarity 83.3%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 2677 CVPITC 2682

RESULT 5
 ID 096RM4 PRELIMINARY; PRT: 3508 AA.
 AC 096RM4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CUB AND SUSHI MULTIPLE DOMAINS 1 PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21365705; PubMed=11472063;
 RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
 RA Sunwoo J.B., Gollin S.M., Scholnick S.B.;
 RT "Transcript map of the 8p23 putative tumor suppressor region."
 RL Genomics 75:17-25(2001).
 DR EMBL: AF333704; AAK73475.1; -
 SO SEQUENCE 3508 AA; 382824 MW; 9268C3BEF3F78C18 CRC64;

Query Match 94.7%; Score 36; DB 4; Length 3508;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 2677 CVPITC 2682

RESULT 6
 ID 0923L3 PRELIMINARY; PRT: 3564 AA.
 AC 0923L3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CSMD1.
 GN CSMD1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;

RX MEDLINE-21365705; PubMed-11472063;
 RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,
 RA Sunwoo J.B., Gollin S.M., Scholnick S.B.;
 RT "transcript map of the 8p23 putative tumor suppressor region.";
 RL Genomics 75:17-25(2001).
 DR EMBL: AY017475; AAC54083.1;
 SO SEQUENCE 3564 AA; 387865 MW; 70824C55B0674609 CRC64;

Query Match 94.7%; Score 36; DB 11; Length 3564;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 2733 CVPITC 2738

RESULT 7
 ID 09DGM2 PRELIMINARY; PRT; 426 AA.
 AC 09DGM2;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EMDOTHELIN RECEPTOR BL.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RX SEQUENCE FROM N.A.
 RA Parichy D.M., Kesh R.N., Mellgren E., Lopes S., Rawls J.,
 RA Johnson S.L.;
 RT "Mutational analysis of roles for endothelin receptor bl during neural
 crest and pigment pattern development in the zebrafish, Danio rerio.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275636; AAC00977.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT VARIANT 156 156 C -> *.
 FT VARIANT 184 184 A -> E.
 FT VARIANT 318 318 W -> *.
 SO SEQUENCE 426 AA; 48770 MW; F92C0B0B3046E4A3 CRC64;

Query Match 92.1%; Score 35; DB 13; Length 426;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 265 CVPITC 270

RESULT 8
 ID 09JUK57 PRELIMINARY; PRT; 468 AA.
 AC 09JUK57;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PREGNANCY-ASSOCIATED PLASMA PROTEIN A (FRAGMENT).
 GN PAPPA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RA Olesen C., Hansen C., Hayashizaki Y., Bykov A., Tommerup N.;
 RT "Partial sequence of Mus musculus pregnancy-associated
 RT plasma protein A (Pappa).";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF260433; AAF70319.1;
 DR MGD: MGI:97479; Pappa.
 DR InterPro: IPR000800; Notch.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00084; sush1. 5.
 DR SMART: SM00032; CCP; 4.
 DR SMART: SM00004; NL; 1.
 FT NON_TER 1
 SO SEQUENCE 468 AA; 51134 MW; 85768E10D7D34FDA CRC64;

Query Match 92.1%; Score 35; DB 11; Length 468;
 Best Local Similarity 83.3%; Pred. No. 11;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 251 CVPITC 256

RESULT 9
 ID 096Q03 PRELIMINARY; PRT; 946 AA.
 AC 096Q03;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE KIA1884 PROTEIN (FRAGMENT).
 GN KIA1884.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RA TISSUE-BRAIN;
 RC MEDLINE-21456161; PubMed-11572484;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XXI.
 RT The complete sequences of 60 new cDNA clones from brain which code for
 RT large proteins.";
 RL DNA Res. 8:179-187(2001).
 DR EMBL: AB067471; BAB67777.1;
 DR NON_TER 1
 SO SEQUENCE 946 AA; 102543 MW; EDC903C6BD390B4A CRC64;

Query Match 92.1%; Score 35; DB 4; Length 946;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPITC 6
 DB 47 CVPITC 52

RESULT 10
 ID 09ES06 PRELIMINARY; PRT; 1367 AA.
 AC 09ES06;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PREGNANCY-ASSOCIATED PLASMA PROTEIN-A (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Hourvitz A., Kuwahara A., Hennebold J.D., Negishi H., Tanaka M.,
 Wildger A., Erickson G.F., Adashi E.Y.;
 RT "The regulated expression of insulin-like growth factor binding
 protein-4-endopeptidase (Pregnancy-associated plasma protein-A) in the
 rodent ovary: A marker of the dominant follicle and of the corpus
 luteum.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF258461; AAG09799.1;
 DR InterPro: IPR000800; Notch.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR000130; Zn_MTpeptidase.
 DR Pfam: PF00084; Sushi. 5.
 DR SMART: SM00032; CCP. 4.
 DR SMART: SM00004; NL; 3.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 FT NON_TER 1
 SO SEQUENCE 1367 AA; 152644 MW; A46C5B049042A5 CRC64;

Query Match 92.1%; Score 35; DB 11; Length 1367;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
 Db 1150 CVPVTC 1155

RESULT 11
 O9XU36 PRELIMINARY; PRT; 367 AA.
 AC O9XU36;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE F1AF8.5 PROTEIN.
 GN F1AF8.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd C.R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z92782; CAB07186.1;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR PRINTS: PR00237; GPCRHHODOPSN.
 DR PROSITE: PS50362; G_PROTEIN_RECIP_F1_2; 1.
 SO SEQUENCE 367 AA; 42394 MW; 4C72196AEAEF1F8D CRC64;

Query Match 89.5%; Score 34; DB 5; Length 367;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
 Db 92 CVPVTC 97

RESULT 12
 O9PTM2 PRELIMINARY; PRT; 66 AA.
 AC O9PTM2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE JAC.
 GN JAC.
 OS Coturnix coturnix (common quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=9091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20238035; PubMed=10772983;
 RA Bader A.G., Hartl M., Bister K.;
 RT "Conditional cell transformation by doxycycline-controlled expression
 of the ASV17 v-jun allele.";
 RL Virology 270:98-110(2000).
 DR EMBL: AF172322; AAD51654.1;
 SO SEQUENCE 66 AA; 6955 MW; 99539EC35A046B89 CRC64;

Query Match 86.8%; Score 33; DB 13; Length 66;
 Best Local Similarity 66.7%; Pred. No. 5.4;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
 Db 43 CVPVTC 48

RESULT 13
 O9AIP4 PRELIMINARY; PRT; 66 AA.
 AC O9AIP4;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN SPY0169.
 GN SPY0169.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillales/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Ian S.P.,
 Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL: AE006486; AAK33269.1;
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 66 AA; 7347 MW; 73BA4A2AA2538C5C CRC64;

Query Match 86.8%; Score 33; DB 16; Length 66;
 Best Local Similarity 83.3%; Pred. No. 5.4;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVPLTC 6
 Db 45 CVPVTC 50

RESULT 14
 O9PTM3

ID 09PTM3 PRELIMINARY; PRT; 68 AA.
 AC 09PTM3;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE JAC (JAC PROTEIN).
 GN JAC.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bader A.G., Hartl M., Bister K.;
 RT "Conditional cell transformation by doxycycline-controlled expression
 of the ASV17 v-jun allele.";
 RL Virology 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20238035; PubMed-10772983;
 RA Bader A.G., Hartl M., Bister K.;
 RT "Conditional cell transformation by doxycycline-controlled expression
 of the ASV17 v-jun allele.";
 RL Virology 270:98-110(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Hartl M., Reiter F., Bader A.G., Bister K.;
 RT "Structure of the JAC gene locus, a putative AP-1 target.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF172321; AAD51653.1;
 DR EMBL; AF239161; AAF42955.1;
 SQ SEQUENCE 68 AA; 7044 MW; 084ED68EBCA1F1B7 CRC64;

Query Match 86.8%; Score 33; DB 13; Length 68;
 Best Local Similarity 66.7%; Pred. No. 5.6;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVPLTC 6
 Db 45 CMPPTC 50

RESULT 15
 078161
 ID 078161 PRELIMINARY; PRT; 93 AA.
 AC 078161;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ENVELOPE PROTEIN (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-PATIENT RI;
 RC Sala M., Zambreno G., Vartanian J., Marconi A., Bertazzoni U.,
 RA Wain-Hobson S.;
 RT "Spatial discontinuities in human immunodeficiency virus type 1
 quasispecies derived from epidermal langerhans cells of an AIDS
 patient and evidence for double infection.";
 RL J. Virol. 0:0-0(1994).
 DR EMBL; Z34348; CAA84119.1;
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 93
 SQ SEQUENCE 93 AA; 10300 MW; B8FFD80DD35FE626 CRC64;

Query Match 86.8%; Score 33; DB 15; Length 93;
 Best Local Similarity 83.3%; Pred. No. 7.3;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CVPLTC 6
 Db 6 CVPLNC 11

Search completed: June 17, 2002, 16:23:44
 Job time: 1618 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:15 ; Search time 227.35 Seconds
(without alignments)
2.931 Million cell updates/sec

Title: US-09-761-636a-12

Perfect score: 38
Sequence: 1 CVPRLRC 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_032802:*

- 1: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	6	22	AAU04531
2	36	94.7	613	22	AAU00816
3	36	94.7	882	22	AAU00816
4	36	94.7	883	22	AAU00816
5	35	92.1	95	22	AAU04819
6	35	92.1	1603	22	AAU03249
7	35	92.1	1752	22	ABG11138
8	33	86.8	10	21	AAU0081
9	33	86.8	60	22	AAU06225
10	33	86.8	60	22	AAU06225
11	33	86.8	138	22	AAU067640

12	33	86.8	194	22	AAU55878	Proionibacterium
13	33	86.8	637	20	AAU25004	Melon MEERS proteol
14	33	86.8	688	9	AAU0361	Tick WGT antigen
15	33	86.8	1236	22	AAU0480	Drosophila melanog
16	33	86.8	1524	20	AAU15458	Human laminin gann
17	33	86.8	1587	21	AAU04917	Human ORFX ORF681
18	32	84.2	49	21	AAU91362	Human secreted pro
19	32	84.2	111	22	AAU03492	Human polypeptide
20	32	84.2	125	22	AAU31573	Peptide #4224 enco
21	32	84.2	125	22	AAU17763	Peptide #4197 enco
22	32	84.2	125	22	AAU30275	Peptide #4312 enco
23	32	84.2	153	21	AAU28004	Arabidopsis thalia
24	32	84.2	162	21	AAU28004	Arabidopsis thalia
25	32	84.2	235	22	AAU75526	Human colon cancer
26	32	84.2	262	22	AAU04267	Human gene 8 enco
27	32	84.2	289	22	AAU16507	Novel human diagn
28	32	84.2	361	22	AAU99191	Human polypeptide
29	32	84.2	409	22	AAU04531	Human human diagn
30	32	84.2	727	11	AAU05533	Fragment of Heyman
31	32	84.2	742	22	AAU59085	Drosophila melanog
32	32	84.2	1761	20	AAU15457	Human laminin beta
33	32	84.2	3396	22	AAU64261	Drosophila melanog
34	32	84.2	3478	22	AAU04526	Novel human diagn
35	32	84.2	4655	17	AAU97208	Human calcium sens
36	32	84.2	4655	17	AAU97209	Human calcium sens
37	32	84.2	4655	17	AAU97210	Human kidney calci
38	32	84.2	4655	17	AAU97211	Human parathyroid
39	32	84.2	4655	19	AAU43311	Human calcium sens
40	32	84.2	4655	19	AAU43312	Human placental ca
41	32	84.2	4655	19	AAU43313	Human kidney calci
42	32	84.2	4655	19	AAU43314	Human parathyroid
43	32	84.2	4689	22	AAU04530	Novel human diagn
44	31	81.6	55	22	AAU24099	Purple urchin EST
45	31	81.6	63	22	AAU34466	Peptide #1972 enco

ALIGNMENTS

RESULT 1	
AAU04531	standard: Peptide; 6 AA.
XX	
XX	AAU04531;
XX	
XX	26-SEP-2001 (first entry)
XX	
XX	VEGF based monocyclic peptide 9.
DE	
XX	Human; VEGF: vascular endothelial growth factor; angiogenesis;
KW	neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW	diabetes induced neovascular sequelae; rheumatoid arthritis;
KW	diabetic retinopathy; chronic inflammation; cyclic.
OS	Synthetic.
XX	
XX	Key Location/Qualifiers
FT	Disulfide-bond 1..6
FT	/note="this bond cyclises the peptide"
XX	
XX	WO200152875-A1.
XX	
XX	26-JUL-2001.
XX	
XX	18-JAN-2001; 2001WO-US01533.
XX	
XX	18-JAN-2000; 2000US-0176293.
XX	
XX	16-MAY-2000; 2000US-0204590.
XX	
XX	(LUDW-) LUDWIG INST CANCER RES.
XX	
XX	Achen MG, Hughes RA, Stacker S, Cendron A;
XX	

DR WPI: 2001-442248/47.

XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
XX oxidizing the cysteine residues -

PS Claim 49: Page 32; 102pp: English.

XX The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the exposed loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.

XX Sequence 6 AA:

Query Match 100.0%; Score 38; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CYPITC 6
Db 1 CYPITC 6

RESULT 2

ID AAU00816 standard: Protein; 613 AA.

AC AAU00816;

DT 01-JUN-2001 (first entry)

XX Human Immunoglobulin superfamily, IgSF, protein #2.

XX Human; Immunoglobulin superfamily protein; IgSF; Immune response;
KW inflammatory response; cell-cell interaction; cell-surface recognition;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; renal disorder; proliferative disorder;
KW cancer; common variable adhesion deficiency syndrome; AIDS; SCID;
KW Alzheimer's disease; Crohn's disease; anaemia; rheumatoid arthritis;
KW Hodgkin's lymphoma; Parkinson's disease; liver cancer; leukaemia;
KW arteriosclerosis; stroke; diabetes mellitus; Addison's disease;
KW urticaria; severe combined immunodeficiency; antibody.

OS Homo sapiens.

XX

FH Key Location/Qualifiers
FT Domain 4..52
FT /label= Immunoglobulin-like-domain
FT Region 6..11
FT /label= Immunogenic-epitope
FT Region 73..79
FT /label= Immunogenic-epitope
FT Region 121..127
FT /label= Immunogenic-epitope
FT Region 141..146
FT /label= Immunogenic-epitope
FT Region 223..230
FT /label= Immunogenic-epitope
FT Region 253..260
FT /label= Immunogenic-epitope
FT Region 328..336
FT /label= Immunogenic-epitope
FT Region 374..383
FT /label= Immunogenic-epitope
FT Region 392..404
FT /label= Immunogenic-epitope
FT Region 484..500
FT /label= Immunogenic-epitope
FT Region 511..517
FT /label= Immunogenic-epitope
FT Region 577..583
FT /label= Immunogenic-epitope
FT Region 605..610
FT /label= Immunogenic-epitope
XX WO200118176-A1.
XX 15-MAR-2001.
XX 29-AUG-2000; 2000WO-US23662.
XX 03-SEP-1999; 99US-0152248.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (NIJ/) NI J.
XX Young PE, Ruben SM, Shi Y;
XX WPI: 2001-203084/20.
XX N-PSDB; AAS00146.
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -
XX Claim 11: Page 238-239; 247pp: English.

XX The sequence is a Human Immunoglobulin superfamily, IgSF, protein,
CC a diverse family of proteins involved in cell-cell interactions,
CC cell-surface recognition, intercellular communication and immune and
CC inflammatory responses. Polypeptides and antibodies directed to
CC polypeptides of the present invention are useful to provide immunological
CC probes for differential identification of tissues. Antibodies can be
CC used to assay levels of polypeptides encoded by polynucleotides of the
CC invention. Polypeptides of the present invention can be used to treat
CC or prevent diseases or conditions such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular
CC disorders, renal disorders, proliferative disorders, and/or cancerous
CC diseases and conditions. Polynucleotides of the invention are also
CC useful in treating the above disorders. Examples of the disorders
CC include common variable adhesion deficiency syndrome, acquired
CC immunodeficiency syndrome (AIDS), anaemia, rheumatoid arthritis,
CC Alzheimer's disease, Crohn's disease, liver cancer, leukaemia,
CC lymphoma, Parkinson's disease, Huntington's disease, dementia,
CC arteriosclerosis, stroke, diabetes mellitus, Addison's disease,
CC urticaria, severe combined immunodeficiency (SCID). Many more
CC examples of diseases and disorders are given in the specification.

SQ Sequence 613 AA:

Query Match 94.7%; Score 36; DB 22; Length 613;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVP LTC 6
111:11
Db 264 cvp ltc 269

RESULT 3
ID AAB83372 standard; Protein; 882 AA.
AC AAB83372;
XX 26-MAR-2002 (first entry)
DE NOV16 protein sequence.
XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.
XX Unidentified.
OS WO200136638-A2.
XX 25-MAY-2001.
PD 17-NOV-2000; 2000WO-US31543.
PF 19-NOV-1999; 99US-166336P.
XX 29-NOV-1999; 99US-167785P.
PR 08-MAR-2000; 2000US-187844P.
PR 16-NOV-2000; 2000US-0715417.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Lichenstein H, Vernet C, Fernandes E;
XX WPI; 2001-648134/74.
DR N-PSDB; AAF87127.
XX Novel human polypeptides and the nucleic acids that encode them useful
PT for preventing, diagnosing and treating e.g. cancer, inflammation and
PT immune disorders -
XX Claim 1; Page 50-52; 141pp; English.

This sequence is the NOV16 protein. The invention relates to
CC the NOV1-NOV16 proteins, and their coding sequences. The proteins have
CC Cytostatic; contraceptive; antiinflammatory; immunomodulatory; and
CC cardiovascular activities. The sequences may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate NOVX
CC expression. They may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of protein by expressing inactive proteins or to
CC supplement the patients own production of protein. They are used to
CC produce NOVX proteins, by inserting the nucleic acid into a cell and
CC culturing it to express the protein. The DNA may be used as DNA probes in
CC assays to detect and quantitate the presence of similar DNAs in samples,
CC and which patients may need restorative therapy. The NOVX protein may
CC also be used as antigens in the production of antibodies (Abs) against
CC NOVX and in assays to identify modulators of NOVX expression and

CC activity. The anti-NOVX Abs and antagonist are used to down regulate
CC expression and activity. The anti-NOVX Abs are used for detecting the
CC presence of NOVX in samples. Disorders that may be prevented, diagnosed
CC and/or treated vary depending on the NOVX protein. NOV1, NOV3, NOV5,
CC NOV7, NOV9-11 and NOV13-16 have casein kinase II phosphorylation sites
CC characteristic of serine/threonine kinases, and are used to treat
CC kinase-related disorders (e.g. Peutz-Jeghers syndrome, cellular
CC proliferation and contraception). NOV2-3, NOV6 and NOV8 are homologous to
CC the epidermal growth factor (EGF)-like super family and are involved in,
CC e.g. regulation of cell development, apoptosis, cell adhesion, growth
CC migration, cell structure and motility and protein management, and are
CC used to treat cancers, inflammatory disorders, immune disorders and
CC cellular adhesion disorders. NOV6-10 are homologous to EGF-like
CC fibrillin proteins and are used to treat cardiovascular disease e.g.
CC hypertrophic cardiomyopathy, long-QT syndrome and marfan syndrome.

SQ Sequence 882 AA:

Query Match 94.7%; Score 36; DB 22; Length 882;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CVP LTC 6
111:11
Db 264 cvp ltc 269

RESULT 4
ID AAB83371 standard; Protein; 883 AA.
AC AAB83371;
XX 26-MAR-2002 (first entry)
DE NOV15 protein sequence.
XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.
XX Unidentified.
OS WO200136638-A2.
XX 25-MAY-2001.
PD 17-NOV-2000; 2000WO-US31543.
PF 19-NOV-1999; 99US-166336P.
XX 29-NOV-1999; 99US-167785P.
PR 08-MAR-2000; 2000US-187844P.
PR 16-NOV-2000; 2000US-0715417.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Lichenstein H, Vernet C, Fernandes E;
XX WPI; 2001-648134/74.
DR N-PSDB; AAF87126.
XX Novel human polypeptides and the nucleic acids that encode them useful
PT for preventing, diagnosing and treating e.g. cancer, inflammation and
PT immune disorders -
XX Claim 1; Page 46-48; 141pp; English.

XX WPI: 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
XX Claim 20: Page 624; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 1603 AA;

XX

Sequence 1603 AA;

Query Match 92.1%; Score 35; DB 22; Length 1603;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||:|
DB 1385 cvptc 1390

RESULT 7
ID AAB11138 standard; Protein; 1752 AA.
XX
AC AAB11138;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11129.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS75325.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20; SEQ ID No 41497; 103pp; English.
XX

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1752 AA;

Query Match 92.1%; Score 35; DB 22; Length 1752;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVPLTC 6
|||:|
DB 1483 cvptc 1488

RESULT 8
ID AAB23798 standard; peptide; 10 AA.
XX
AC AAB23798;
XX
DT 15-JAN-2001 (first entry)
XX
DE Phage display peptide K13 SEQ ID NO:1.
XX
KW Phage display peptide; screening; biomolecule activity regulator;
KW drug development; liver disorder; sclerostis; cancer; hepatitis C virus;
KW HCV; infection.
XX
OS Unidentified bacteriophage.
XX
PN WO200053740-A1.
PD 14-SEP-2000.
XX
PF 10-MAR-2000; 2000WO-JP01478.
XX
PR 10-MAR-1999; 99JP-0063110.
XX
PA (AJIN) AJINOMOTO CO INC.
PI Okamoto S, Miwa K, Eto Y;
XX
DR WPI: 2000-587433/55.
XX
PT Screening biomolecule activity regulators by their effect on a
PT biomolecule-peptide interactions for identification of potential drug
PT molecules -
XX
PS Example 2; Page 24; 45pp; Japanese.
XX

CC The present invention describes a method for screening potential
CC regulators of the activity of a biomolecule. The method involves
CC screening the potential regulators by selecting from a library of
CC recombinants which express peptides at their surface, one which
CC expresses a peptide which interacts with the biomolecule and then
CC screening potential regulators to identify those which inhibit the
CC interaction of the biomolecule with the selected recombinant or the
CC peptide expressed by it. The method is used for the identification
CC of drug molecules for treatment and prevention of diseases, especially
CC of liver disorders such as sclerosis and cancer associated with
CC hepatitis C virus (HCV) infection. The present sequence represents a
CC phase display peptide which is used in an example from the present
CC invention.
SQ Sequence 10 AA;

Query Match 86.8%; Score 33; DB 21; Length 10;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 1 cvplvc 6

RESULT 9
AAU40081
ID AAU40081 standard; Protein; 60 AA.
XX
AC AAU40081;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #977.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PE 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB: AAS59510.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 1276; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 60 AA;

Query Match 86.8%; Score 33; DB 22; Length 60;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 28 cvpltc 33

RESULT 10
AAU66225
ID AAU66225 standard; Protein; 60 AA.
XX
AC AAU66225;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #27121.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PE 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB: AAS59722.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 27420; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC XX
SQ Sequence 60 AA;

Query Match 86.8%; Score 33; DB 22; Length 60;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CVPLTC 6
||| |
Db 28 cvpltc 33

RESULT 11
AAU67640
ID AAU67640 standard; Protein: 138 AA.
XX
AC AAU67640;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #26536.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR MPI: 2001-616774/71.
DR N-PSDB; AAS59609.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 28835; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC XX
SQ Sequence 138 AA;

Query Match 86.8%; Score 33; DB 22; Length 138;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CVPLTC 6
||| |
Db 8 cvplac 13

RESULT 12
AAU55878
ID AAU55878 standard; Protein: 194 AA.
XX
AC AAU55878;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #16774.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR MPI: 2001-616774/71.
DR N-PSDB; AAS59572.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 17073; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by

P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 194 AA;

Query Match 86.8%; Score 33; DB 22; Length 194;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVP LTC 6
11111
Db 8 cvplac 13

RESULT 13

AA25004
ID AAY25004 standard; Protein; 637 AA.

XX AAY25004;

XX 23-AUG-1999 (first entry)

XX Melon MEERS protein.

XX MEERS; melon; ethylene receptor; plant; low ethylene sensitivity;

XX agricultural product.

XX Cucumis sp.

XX JP1151090-A.

XX 08-JUN-1999.

XX 31-MAR-1998; 98JP-0086214.

XX 17-SEP-1997; 97JP-0251222.

XX (IBAR-) IBARAGI PREFECTURE.

XX WPI; 1999-388480/33.

XX N-PSDB; AAX78216.

XX New plant of low ethylene sensitivity - useful for improving life of agricultural products

XX Claim 1; Page 10-11; 22pp; Japanese.

XX This invention describes a novel ethylene receptor protein and its encoding nucleic acid isolated from melon (Cucumis sp.) designated MEERS. The introduction of the gene of the invention into a plant results in plants of low ethylene sensitivity. The protein of the invention can improve the life of agricultural products.

XX Sequence 637 AA;

Query Match 86.8%; Score 33; DB 20; Length 637;

Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CVP LTC 6
11111
Db 226 cvpylc 231

RESULT 14

AA25004
ID AAP80361 standard; protein; 688 AA.

XX AAP80361;

XX 03-DEC-1990 (first entry)

XX Tick WGL+ antigen.

XX WGL+ antigen; cattle vaccine.

XX Boophilus microplus.

XX Key Location/Qualifiers

XX Peptide 29..52

XX Peptide 60..83

XX Peptide 82..98

XX Peptide 175..186

XX Peptide 209..220

XX Peptide 222..229

XX Peptide 237..257

XX Peptide 283..291

XX Peptide 367..383

XX Peptide 414..431

XX Peptide 453..460

XX Peptide 497..513

XX Peptide 523..541

XX Peptide 629..635

XX WO8803929-A.

XX 02-JUN-1988.

XX 27-NOV-1987; 87WO-A000401.

XX 16-OCT-1987; 87AU-0004912.

XX 16-OCT-1987; 87AU-0009196.

XX (BIOT-) BIOTECHN AUSTR PTY.

XX (CSIR) COMMONWEALTH SCI IND RES ORG.

XX Cobon GS, Moore JT, Johnston LA, Willadsen P, Kemp DH;

XX Strickland A, Riding GA, Rand KN;

XX WPI; 1988-161607/23.

XX N-PSDB; AAN80329.

XX Vaccine contg. tick antigens - used for immunising a mammalian host,

XX esp. cattle to induce immunity to tick infestation.

XX Claim 19; Page 60; 83pp; English.

XX This tick WGL+ antigen, and peptide fragments of it, is useful as an immunogen to induce immunity to tick infestation in a host species. The resultant immune response produced by the host is capable of damaging the plasma membrane of the gut cells of ticks feeding on the host. It can be used for vaccinating cattle against diseases caused by a wide range of tick species. See also AAN80330-35.

XX Sequence 688 AA;

Query Match 86.8%; Score 33; DB 9; Length 688;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CVP LTC 6
 ||| ||
 Db 157 cvptlc 162

RESULT 15

ABB60480
 ID ABB60480 standard; Protein; 1236 AA.

AC ABB60480:

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 8232.

KW Drosophila: developmental biology; cell signalling; insecticide;
 pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

DR N-PSDB; ABL04583.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 genes from Drosophila and for elucidating cell signalling and cell-cell
 interactions -

PS Disclosure; SEQ ID NO 8232; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 capable of detecting 1000 or more genes from Drosophila. The invention is
 useful in developmental biology and in elucidating cell signalling and
 cell-cell interactions in higher eukaryotes for the development of
 insecticides, therapeutics and pharmaceutical drugs. The invention
 discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 sequences (ABL01840-ABL16175) and the encoded proteins
 (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1236 AA;

Query Match 86.8%; Score 33; DB 22; Length 1236;

Best Local Similarity 83.3%; Pred. NO. 1.8e+03;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6

Db 1196 cvptlc 1201

Search completed: June 17, 2002, 16:02:16
 Job time: 425 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:39 ; Search time 75.98 Seconds

(without alignments)
1.929 Million cell updates/sec

Title: US-09-761-636A-12

Perfect score: 38

Sequence: 1 CVP/TC 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	86.8	549	1	US-08-325-071-61 Sequence 61, Appl
2	33	86.8	549	4	US-08-461-004A-61 Sequence 61, Appl
3	33	86.8	620	1	US-08-325-071-65 Sequence 65, Appl
4	33	86.8	620	4	US-08-461-004A-65 Sequence 65, Appl
5	33	86.8	650	1	US-08-325-071-56 Sequence 56, Appl
6	33	86.8	650	1	US-08-325-071-59 Sequence 59, Appl
7	33	86.8	650	1	US-08-325-071-63 Sequence 63, Appl
8	33	86.8	650	1	US-08-325-071-67 Sequence 67, Appl
9	33	86.8	650	4	US-08-461-004A-56 Sequence 56, Appl
10	33	86.8	650	4	US-08-461-004A-59 Sequence 59, Appl
11	33	86.8	650	4	US-08-461-004A-63 Sequence 63, Appl
12	33	86.8	650	4	US-08-461-004A-67 Sequence 67, Appl
13	33	86.8	688	1	US-08-325-071-57 Sequence 57, Appl
14	33	86.8	688	4	US-08-461-004A-57 Sequence 57, Appl
15	33	84.2	38	6	5208144-23 Patent No. 5208144
16	32	84.2	726	6	5208144-37 Patent No. 5208144
17	32	84.2	4654	4	US-08-476-515A-84 Sequence 84, Appl
18	32	84.2	4655	4	US-08-652-877-84 Sequence 84, Appl
19	32	84.2	4655	4	US-08-652-877-86 Sequence 86, Appl
20	32	84.2	4655	4	US-08-652-877-88 Sequence 88, Appl
21	32	84.2	4655	4	US-08-652-877-90 Sequence 90, Appl
22	31	81.6	39	1	US-08-212-236-9 Sequence 9, Appl
23	31	81.6	40	1	US-08-033-873-5 Sequence 5, Appl
24	31	81.6	40	2	US-08-356-832-5 Sequence 5, Appl
25	31	81.6	40	4	US-08-988-705-5 Sequence 5, Appl
26	31	81.6	810	2	US-08-820-170A-34 Sequence 34, Appl
27	31	81.6	810	3	US-09-055-699-34 Sequence 34, Appl

28	31	81.6	810	4	US-09-273-565-34 Sequence 34, Appl
29	31	81.6	810	4	US-09-565-538-34 Sequence 34, Appl
30	30	78.9	38	1	US-07-603-451A-1 Sequence 1, Appl
31	30	78.9	38	1	US-08-060-832A-1 Sequence 1, Appl
32	30	78.9	38	1	US-08-033-873-8 Sequence 8, Appl
33	30	78.9	38	1	US-08-033-873-15 Sequence 15, Appl
34	30	78.9	38	1	US-08-248-016-10 Sequence 10, Appl
35	30	78.9	38	1	US-08-212-236-6 Sequence 6, Appl
36	30	78.9	38	1	US-08-451-501-10 Sequence 10, Appl
37	30	78.9	38	2	US-08-356-832-8 Sequence 8, Appl
38	30	78.9	38	2	US-08-356-832-15 Sequence 15, Appl
39	30	78.9	38	4	US-08-988-705-8 Sequence 8, Appl
40	30	78.9	38	4	US-08-988-705-15 Sequence 15, Appl
41	30	78.9	38	5	PCT-US94-05257-1 Sequence 1, Appl
42	30	78.9	38	5	PCT-US95-06761-10 Sequence 10, Appl
43	30	78.9	39	1	US-08-212-236-8 Sequence 8, Appl
44	30	78.9	40	1	US-08-033-873-2 Sequence 2, Appl
45	30	78.9	40	1	US-08-033-873-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-325-071-61
; Sequence 61, Application US/08325071
; Patent No. 5587311
GENERAL INFORMATION:
; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
; APPLICANT: JOHNSON, Law Anthony York
; APPLICANT: WILLADSEN, Peter
; APPLICANT: KEMP, David Harold
; APPLICANT: SRISKANTHA, Alagacone
; APPLICANT: RIDING, George Alfred
; APPLICANT: RAND, Keith No. 3587311man
; TITLE OF INVENTION: DNA Encoding A Cell Membrane
; TITLE OF INVENTION: Glycoprotein Of A Tick Gut
; NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,071
; FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,109
; FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,368
; FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/242,196
; FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU87/00401
; FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P14912
; FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P12570
; FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P9196

ATTORNEY/AGENT INFORMATION:
FILING DATE: 27-NOV-1986
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 B1AU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-61

Query Match 86.8%; Score 33; DB 1; Length 549;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6
11111
Db 66 CVP TTC 71

RESULT 2
US-08-461-004A-61
Sequence 61, Application US/08461004A
Patent No. 6235283
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-61

Query Match 86.8%; Score 33; DB 4; Length 549;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6
11111
Db 66 CVP TTC 71

RESULT 3
US-08-325-071-65
Sequence 65, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-65

Query Match 86.8%; Score 33; DB 1; Length 620;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
11111
DB 117 CVPTTC 122

RESULT 4
US-08-461-004A-65
Sequence 65, Application US/08461004A
Patent No. 6235283
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 6235283man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
Protein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109

FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-65

Query Match 86.8%; Score 33; DB 4; Length 620;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
11111
DB 117 CVPTTC 122

RESULT 5
US-08-325-071-56
Sequence 56, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
Protein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P19196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-56

Query Match 86.8%; Score 33; DB 1; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 147 CVPTTC 152

RESULT 6
US-08-325-071-59
Sequence 59, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P19196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-59

Query Match 86.8%; Score 33; DB 1; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
Db 147 CVPTTC 152

RESULT 7
US-08-325-071-63
Sequence 63, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P19196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-63

Query Match 86.8%; Score 33; DB 1; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPFTC 6
Db 147 CVPFTC 152

RESULT 8
US-08-325-071-67
; Sequence 67, Application US/08325071
; Patent No. 5587311
; GENERAL INFORMATION:
; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
; APPLICANT: JOHNSON, Law Anthony York
; APPLICANT: WILLADSEN, Peter
; APPLICANT: KEMP, David Harold
; APPLICANT: SRISKANTHA, Alagacane
; APPLICANT: RIDING, George Alfred
; APPLICANT: RAND, Keith No. 5587311man
; TITLE OF INVENTION: DNA Encoding A Cell Membrane
; TITLE OF INVENTION: Glycoprotein of A Tick Gut
; NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P19196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-071-67

Query Match 86.8%; Score 33; DB 1; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPFTC 6
Db 147 CVPFTC 152

RESULT 9
US-08-461-004A-56
; Sequence 56, Application US/08461004A
; Patent No. 6235283
; GENERAL INFORMATION:
; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
; APPLICANT: JOHNSON, Law Anthony York
; APPLICANT: WILLADSEN, Peter
; APPLICANT: KEMP, David Harold
; APPLICANT: SRISKANTHA, Alagacane
; APPLICANT: RIDING, George Alfred

APPLICANT: RAND, Keith No. 6235283man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI2570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-56

Query Match 86.8%; Score 33; DB 4; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
||| ||
Db 147 CVPTTC 152

RESULT 10
US-08-461-004A-59
; Sequence 59, Application US/08461004A
; Patent No. 6235283
; GENERAL INFORMATION:

APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacoe
APPLICANT: RIDING, George Alfred
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI2570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-59

Query Match 86.8%; Score 33; DB 4; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
||| ||
Db 147 CVPTTC 152

RESULT 11

US-08-461-004A-63

; Sequence 63, Application US/08461004A
; Patent No. 6235283

GENERAL INFORMATION:

; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
; APPLICANT: JOHNSON, Law Anthony York
; APPLICANT: WILLADSEN, Peter
; APPLICANT: KEMP, David Harold
; APPLICANT: SRISKANTHA, Alagacone
; APPLICANT: RIDING, George Alfred
; APPLICANT: RAND, Keith No. 6235283man
; TITLE OF INVENTION: DNA Encoding A Cell Membrane
; TITLE OF INVENTION: Glycoprotein Of A Tick Gut
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,004A
; FILING DATE: 04-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/325,071
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,109
; FILING DATE: 17-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,368
; FILING DATE: 07-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/242,196
; FILING DATE: 06-JUL-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU87/00401
; FILING DATE: 27-NOV-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P14912
; FILING DATE: 16-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P12570
; FILING DATE: 19-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PH9196
; FILING DATE: 27-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 60042/152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136

INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-004A-63Query Match 86.8%; Score 33; DB 4; Length 650;
Best Local Similarity 83.3%; Pred No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6

Db 147 CVP TTC 152

RESULT 12

US-08-461-004A-67

; Sequence 67, Application US/08461004A
; Patent No. 6235283

GENERAL INFORMATION:

; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
; APPLICANT: JOHNSON, Law Anthony York
; APPLICANT: WILLADSEN, Peter
; APPLICANT: KEMP, David Harold
; APPLICANT: SRISKANTHA, Alagacone
; APPLICANT: RIDING, George Alfred
; APPLICANT: RAND, Keith No. 6235283man
; TITLE OF INVENTION: DNA Encoding A Cell Membrane
; TITLE OF INVENTION: Glycoprotein Of A Tick Gut
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,004A
; FILING DATE: 04-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/325,071
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,109
; FILING DATE: 17-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,368
; FILING DATE: 07-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P14912
; FILING DATE: 16-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P12570
; FILING DATE: 19-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PH9196
; FILING DATE: 27-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 60042/152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:

LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-67

Query Match 86.8%; Score 33; DB 4; Length 650;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
11111
Db 147 CVPTTC 152

RESULT 13
US-08-325-071-57
Sequence 57, Application US/08325071
Patent No. 5587311

GENERAL INFORMATION:

APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacane
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI2570
FILING DATE: 19-JUN-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 60042/111 BIAU
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 672 5300

TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ. ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-325-071-57

Query Match 86.8%; Score 33; DB 1; Length 688;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVPLTC 6
11111
Db 157 CVPTTC 162

RESULT 14
US-08-461-004A-57
Sequence 57, Application US/08461004A
Patent No. 6235283

GENERAL INFORMATION:

APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacane
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 6235283man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI2570
FILING DATE: 19-JUN-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-461-004A-37

Query Match 86.8%; Score 33; DB 4; Length 688;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6
11111
DB 157 CVP TTC 162

RESULT 15
5208144-23
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.; RAYCHOMDHURY, RAKTIMA; NILES, JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO: 23:
; LENGTH: 38
5208144-23

Query Match 84.2%; Score 32; DB 6; Length 38;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CVP LTC 6
11111
DB 2 CVP TTC 7

Search completed: June 17, 2002, 15:56:39
Job time: 88 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:26 ; Search time 95.43 seconds
(without alignments)
10.069 Million cell updates/sec

Title: US-09-761-636a-13

Perfect score: 56
Sequence: 1 C1SVPLSVPC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	69.6	85	2 T44555	hypothetical prote
2	39	69.6	148	2 G82599	hypothetical prote
3	38	67.9	502	2 P86298	IMP dehydrogenase
4	38	67.9	503	1 JC4999	IMP dehydrogenase
5	38	67.9	616	2 D11864	hypothetical prote
6	37	66.1	485	2 F82521	hypothetical prote
7	37	66.1	485	2 A33647	sulfated surface g
8	36	64.3	134	2 D45835	Ly6 homolog Rk3 pr
9	36	64.3	207	2 B86781	hypothetical prote
10	36	64.3	256	2 G75515	conserved hypotet
11	36	64.3	496	2 G83136	probable amino aci
12	36	64.3	512	2 H96759	probable steroid 2
13	36	64.3	581	2 S46742	hypothetical prote
14	36	64.3	669	2 T08827	hypothetical prote
15	35	62.5	100	2 F72751	hypothetical prote
16	35	62.5	255	2 JC7593	SH2 domain-contain
17	35	62.5	323	2 I49529	transcription fact
18	35	62.5	372	2 S72711	masc protein - Myc
19	35	62.5	382	2 A63171	probable aminotran
20	35	62.5	438	2 T15039	omega-3 fatty acid
21	35	62.5	468	2 T50982	origin recognition
22	35	62.5	503	2 D75437	DNA repair protein
23	35	62.5	578	2 A70877	probable acyl-coas
24	35	62.5	579	2 B86926	acyl-CoA synthetas
25	35	62.5	579	2 D87063	probable acyl-CoA
26	35	62.5	580	2 B70668	probable acyl-CoA
27	35	62.5	706	2 D71289	probable single-st
28	35	62.5	759	2 G86506	hypothetical prote
29	35	62.5	759	2 G72115	hypothetical prote

30	35	62.5	1162	2 B97852	hypothetical prote
31	34	60.7	38	2 C54471	aglitoxin 3 - scorp
32	34	60.7	77	2 D97718	acetate kinase (Ac
33	34	60.7	156	2 H71023	hypothetical prote
34	34	60.7	249	2 AD0966	probable prs syste
35	34	60.7	250	2 S06314	regulatory protein
36	34	60.7	252	2 B33538	regulatory protein
37	34	60.7	328	2 F89914	hypothetical prote
38	34	60.7	387	2 T44873	probable secreted
39	34	60.7	399	2 AD2559	hypothetical prote
40	34	60.7	410	2 I38502	gene Bm-3b protei
41	34	60.7	440	2 I48291	transcription fact
42	34	60.7	461	1 KXH0	protein C (activat
43	34	60.7	502	2 C95349	hypothetical prote
44	34	60.7	527	2 H85135	hypothetical prote
45	34	60.7	554	2 JEO303	propenadhol dehydr

ALIGNMENTS

RESULT 1
T44555
hypothetical protein PA0631 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 06-Oct-2000
C:Accession: T44555; E83568
R:Nakayama, K.; Takashima, K.; Ishihara, H.; Shinomiya, T.; Kagayama, M.; Kanaya, submitted to the EMBL Data Library, August 1999
A:Description: Genetic relationship between bacteriocins and bacteriophages.
A:Reference number: Z22790
A:Accession: T44555
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-85 <NAK>
A:Cross-references: EMBL:AB030825; PIDN:BAA83170.1
A:Experimental source: strain PA01
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: AB2950; MUID:20437337
A:Accession: E83568
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-85 <STO>
A:Cross-references: GB:AE004498; GB:AE004091; NID:g9946491; PIDN:AGC04020.1; GSPD:AE004498
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0631

Query Match 69.6%; Score 39; DB 2; Length 85;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Gy 2 ISVPLSVPC 10
Db 27 VEYPLAVPC 35

RESULT 2
G82599
hypothetical protein XE2118 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82599
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MUID:20365717
A:Note: For a complete list of authors see reference number A59328 below

A:Accession: G82599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <SIM>
A:Cross-references: GB:AE004026; GB:AE003849; NID:9107228; PIDN:AAF84917.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Bilones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Fachinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Lajth
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuboko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2118

Query Match 69.6%; Score 39; DB 2; Length 148;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 C1SVPLSPC 10
1 :||:|
Db 8 CCAVPMNAPC 17

RESULT 3

F86298
IMP dehydrogenase (EC 1.1.1.205) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: F86298
R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: F86298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <STO>
A:Cross-references: GB:AE005172; NID:94966356; PIDN:AAD34687.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Keywords: GMP biosynthesis; NAD; oxidoreductase

Query Match 67.9%; Score 38; DB 2; Length 502;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 VPLSVPC 10
||||:|
Db 49 VPLSIPC 55

RESULT 4

JC4999
IMP dehydrogenase (EC 1.1.1.205) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JC4999
R:Collart, F.R.; Ositluk, J.; Trent, J.; Olsen, G.J.; Huberman, E.
Gene 174, 217-220, 1996
A:Title: Cloning and characterization of the gene encoding IMP dehydrogenase from
A:Reference number: JC4999; MUID:97045815
A:Accession: JC4999
A:Molecule type: DNA
A:Residues: 1-503 <COL>
A:Cross-references: GB:L34684; NID:91100062; PIDN:AA641940.1; PID:91100063
C:Genetics:
A:Gene: Impdh
A:Insertions: 135/1; 334/3; 404/3; 490/3
C:Complex: homotrimer
C:Function:
A:Description: provides precursors for DNA and RNA biosynthesis. It catalyzes the
C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal h
C:Keywords: GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosynthesis
F:172-220/Domain: CBS homology <CBS2>
F:172-86/Domain: IMP dehydrogenase amino-terminal homology <IDHN>
F:221-465/Domain: IMP dehydrogenase catalytic homology <IMP>
F:322/Active site: Cys #status predicted

Query Match 67.9%; Score 38; DB 1; Length 503;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 VPLSVPC 10
||||:|
Db 49 VPLSIPC 55

RESULT 5

D71864
hypothetical protein jhp0985 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: D71864
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith,
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis,
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastri
A:Reference number: A71800; MUID:9120557
A:Accession: D71864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-616 <ARN>
A:Cross-references: GB:AE001527; GB:AE001439; NID:94155558; PIDN:AAD6568.1; PID:g
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0985

Query Match 67.9%; Score 38; DB 2; Length 616;
Best Local Similarity 77.8%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 C1SVPLSPV 9
||:|||||
Db 253 C1SNPLSPV 261

RESULT 6

F82521
hypothetical protein XP7736 (imported) - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82521
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <STM>
A:Cross-references: GB:AE004080; GB:AE003849; NID:g9107971; PIDN:AAF85521.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brielson, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sasaki
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2736

Query Match 66.1%; Score 37; DB 2; Length 89;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 ISVPLSVP 10
DB 19 VSAPLRVPC 27

RESULT 7
A33647
sulfated surface glycoprotein 185 - Volvox carteri
C:Species: Volvox carteri
C:Date: 11-Apr-1990 #sequence_revision 11-Apr-1990 #text_change 21-Jul-2000
C:Accession: A33647
R:Erli, H.; Mengele, R.; Wenzl, S.; Engel, J.; Sumper, M.
J. Cell Biol. 109, 3493-3501, 1989
A:Title: The extracellular matrix of Volvox carteri: molecular structure of the cellular
A:Reference number: A33647; MUID:90094551
A:Accession: A33647
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <ERT>
A:Cross-references: GB:X51616; NID:g21999; PIDN:CAA35953.1; PID:g1405821
C:Keywords: glycoprotein

Query Match 66.1%; Score 37; DB 2; Length 485;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLSVP 10
DB 83 CQVPLRPEC 92

RESULT 8
D45835
Ly6 homolog RK3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 31-Jan-2000
C:Accession: D45835; C45835
R:Friedman, S.; Paltee, R.G.E.; Stirlin, S.; Haemmerling, U.
Immunogenetics 31, 104-111, 1990
A:Title: Analysis of three distinct Ly6-A-related cDNA sequences isolated from rat kidney
A:Reference number: A45835; MUID:90152758
A:Accession: D45835

A:Molecule type: mRNA
A:Residues: 1-134 <FR1>
A:Cross-references: GB:M30690; NID:g205249; PIDN:AAA41547.1; PID:g205250
A:Experimental source: clone RK3
A:Accession: C45835
A:Molecule type: mRNA
A:Residues: 2-134 <FR2>

A:Cross-references: GB:M30691; NID:g205251; PIDN:AAA41548.1; PID:g205252
A:Experimental source: clone RK11
C:Superfamily: Ly-6 antigen; Ly-6 homolog
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol
F:1-26/Domian: signal sequence #status predicted <SIG>
F:27-105/Product: Ly6 homolog RK3 #status predicted <MAT>
F:106-134/Domian: carboxyl-terminal propeptide #status predicted <CTP>
F:105/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in matu

Query Match 64.3%; Score 36; DB 2; Length 134;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISVPLSVP 10
DB 32 CIEVPLNANC 41

RESULT 9
B86781
hypothetical protein ymhc [imported] - Lactococcus lactis subsp. lactis (strain II
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: B86781
R:Boletín, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lac
A:Reference number: B86625; MUID:21235186; PMID:11337471
A:Accession: B86781
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <STO>
A:Cross-references: GB:AE005176; PID:g12724223; PIDN:AAK05348.1; GSPDB:GN00146
A:Experimental source: strain ILL403
C:Genetics:
A:Gene: ymhc

Query Match 64.3%; Score 36; DB 2; Length 207;
Best Local Similarity 55.6%; Pred. No. 48;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 CISVPLSVP 9
DB 128 CINIPLQIP 136

RESULT 10
G75515
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: G75515
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson,
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski,
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans
A:Reference number: A75250; MUID:20036896
A:Accession: G75515
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <NHT>
A:Cross-references: GB:AE001906; GB:AE000513; NID:g6458151; PIDN:AAF10050.1; PID:g
A:Experimental source: strain R1

C:Genetics:
A:Gene: DR0471
A:Map position: 1

Query Match 64.3%; Score 36; DB 2; Length 256;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 C1SVPLSVPC 10
DB 4 CVPAPRSAPC 13

RESULT 11
G83136
probable amino acid permease PA4072 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83136
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Braham, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Ladbeg, K.; Lim, L.; Lory, S.; Olson, M.V.
Nucleotide 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: G83136
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-486 <STO>
A:Cross-references: GB:AE004823; GB:AE004091; NID:g9950265; PIDN:AAG07459.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4072

Query Match 64.3%; Score 36; DB 2; Length 496;
Best Local Similarity 50.0%; Pred. No. 11e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 C1SVPLSVPC 10
DB 345 CIAVANAIPC 354

RESULT 12
H96759
probable steroid 22-alpha-hydroxylase T9L24.44 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96759
R:Thelander, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.M.; Chou, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Jensen, N.E.; Hughes, B.; Huizar, L.
Nucleotide 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: H96759
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005173; NID:g11120803; PIDN:AAG30983.1; GSPDB:GN00141
C:Genetics:
A:Gene: T9L24.44
A:Map position: 1

Query Match 64.3%; Score 36; DB 2; Length 512;

Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVPLSVPC 9
DB 244 C1SVPLSVPC 252

RESULT 13
S46742
hypothetical protein YHR032w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein H8179.15
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-2001
C:Accession: S46742
R:Du, Z.
Submitted to the EMBL Data Library, May 1994
A:Description: The sequence of S. cerevisiae cosmid 8179.
A:Reference number: S46732
A:Accession: S46742
A:Molecule type: DNA
A:Residues: 1-581 <DU>
A:Cross-references: EMBL:U00062; NID:g4488162; PIDN:AAB68911.1; PID:g4488176; MIPS:Y
C:Genetics:
A:Map position: 8R
C:Keywords: transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 581;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 C1SVPLSVPC 10
DB 564 C1SVPLSVPC 573

RESULT 14
T08827
hypothetical protein cotel - human
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T08827
R:Winfield, S.L.; Tayebi, N.; Martin, B.M.; Gims, E.I.; Sidransky, E.
Genome Res. 7, 1020-1026, 1997
A:Title: Identification of three additional genes contiguous to the glucocerebrosidase
A:Reference number: Z16482; MUID:97474796
A:Accession: T08827
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-669 <WIN>
A:Cross-references: EMBL:AF023268; NID:g2564910; PIDN:AAC51822.1; PID:g2564916
C:Genetics:
A:Gene: cotel
A:Map position: 1
A:Introns: 75/3; 94/3; 131/3; 171/3; 207/3; 266/2; 299/3; 323/1; 505/1; 528/1; 612/1

Query Match 64.3%; Score 36; DB 2; Length 669;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 C1SVPLSVPC 10
DB 139 C1SVPLSVPC 148

RESULT 15
F72751
hypothetical protein APE0003 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: F72751

R;KawadaYasui, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A;Reference number: A72450; MUID:99310339
A;Accession: F72751
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-100 <KAW>
A;Cross-references: DDBJ:AP000058; NID:95103388; PIDN:BAA/8912.1; PID:95103391
A;Experimental source: strain KI
C;Genetics:
A;Gene: APE0003
C;Superfamily: Aeropyrum pernix hypothetical protein APE0003

Query Match 62.5%; Score 35; DB 2; Length 100;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 4 VPLSVPC 10
:|:|:
Db 57 IPLTVPC 63

Search completed: June 17, 2002, 15:58:27
Job time: 196 sec

f

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:31 ; Search time 44.84 Seconds
(without alignments)
8.635 Million cell updates/sec

Title: US-09-761-636a-13
Perfect score: 56
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	75.0	569	1	MANC_PIRSP
2	42	75.0	571	1	MANB_PIRSP
3	38	67.9	502	1	IMH2_ARATH
4	38	67.9	503	1	IMH1_ARATH
5	37	66.1	616	1	Y396_HELPJ
6	37	66.1	420	1	CG2A_CHLVR
7	37	66.1	485	1	SSGP_VOLCA
8	36	64.3	581	1	YHJ2_YEAST
9	36	64.3	669	1	CORE_HUMAN
10	35	62.5	298	1	YMA2_MYCBO
11	35	62.5	323	1	MAF1_MOUSE
12	35	62.5	323	1	MAF1_MOUSE
13	35	62.5	369	1	Y181_FOMPV
14	35	62.5	578	1	FD21_MYCTU
15	35	62.5	579	1	FD21_MYCTU
16	34	61.6	120	1	SY23_HUMAN
17	34	60.7	38	1	SCA3_LEIOH
18	34	60.7	250	1	LUXR_VIBRI
19	34	60.7	336	1	V027_FOMPV
20	34	60.7	410	1	BR3B_HUMAN
21	34	60.7	433	1	RTCL_DICDI
22	34	60.7	461	1	PRTC_HUMAN
23	34	60.7	354	1	DHAB_SALTY
24	34	60.7	590	1	MP44_SFVKA
25	34	60.7	606	1	MANA_PIRSP
26	34	60.7	829	1	CADG_HUMAN
27	33	58.9	348	1	OPSB_HUMAN
28	33	58.9	369	1	OPSB_SALTB
29	33	58.9	369	1	MAF2_RAT
30	33	58.9	369	1	TMAF_AVIS4
31	33	58.9	422	1	MAF2_MOUSE
32	33	58.9	470	1	Y140_HUMAN
33	33	58.9	465	1	YGT4_YEAST

34	33	58.9	573	1	C114_MOUSE
35	33	58.9	583	1	FD25_MYCTU
36	33	58.9	662	1	T982_MOUSE
37	33	58.9	663	1	T982_HUMAN
38	33	58.9	663	1	T982_HUMAN
39	32	57.1	61	1	IBPL_TACTR
40	32	57.1	84	1	SCXA_MESMA
41	32	57.1	210	1	GTT2_MUSDO
42	32	57.1	210	1	GTT3_MUSDO
43	32	57.1	210	1	GTT4_MUSDO
44	32	57.1	219	1	REP2_SCHPO
45	32	57.1	233	1	ATPL_OCHNE
			244	1	PLFR_MOUSE

ALIGNMENTS

RESULT ID	MANC_PIRSP	STANDARD	PRT	569 AA.
AC	P55298			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Mannan endo-1,4-beta-mannosidase C precursor (EC 3.2.1.78) (Beta-mannanase C) (1,4-beta-D-mannan mannanohydrolase C).			
GN	MANC.			
OS	Pirmyces sp.			
OC	Eukaryota; Fungi; Chytridiomycota; Neocallimastixales;			
OC	Neocallimastixaceae; Pirmyces.			
OX	NCBI_TaxID=45796;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9631314; PubMed=8768520;			
RA	Millward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P.,			
RT	Gilbert H.J.;			
RT	"Evidence that the Pirmyces gene family encoding endo-1,4-mannanases arose through gene duplication."			
RL	FEMS Microbiol. Lett. 141:183-188(1996).			
CC	-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.			
CC	-1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANA, MANC AND XYNA; AND TO THOSE OF N.PATRICIARUM XYNA.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
DR	EMBL; X97520; CAA6134.1; -			
DR	InterPro; IPR002883; CBD_5.			
DR	InterPro; IPR008085; Glyco_hydro_26.			
DR	Pfam; PF02013; CBD_5; 2.			
DR	Pfam; PF02156; Glyco_hydro_26; 1.			
DR	PRINTS; PR00739; GLYHDRLASE26.			
KW	Hydrolase; Glycosidase; Signal; Multigene family; Repeat.			
FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	569	MANNAN ENDO-1,4-BETA-MANNOSIDASE C.
FT	DOMAIN	490	569	2 x 39 AA APPROXIMATE REPEATS.
FT	REPEAT	488	525	1.
FT	REPEAT	531	569	2.
FT	DOMAIN	473	477	POLY-ASN.
FT	DOMAIN	480	486	POLY-ASN.
SO	SEQUENCE	569 AA;	64115 MW;	19277764E18328B5 CRC64;

Query Match 75.0%; Score 42; DB 1; Length 569;
Best Local Similarity 60.0%; Pred. No. 4.3;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLSVPC 10
| | | | |
Db 490 CFSIPLGYPC 499

RESULT 2
MANB_PIRSP STANDARD; PRT: 571 AA.

ID MANB_PIRSP STANDARD; PRT: 571 AA.

AC P55297.

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Mannan endo-1,4-beta-mannosidase B precursor (EC 3.2.1.78) (Beta-mannanase B) (1,4-beta-D-mannan mannanhydrolase B).

CN MANB.

OS *Pyromyces* sp.

OC Eukaryota; Fungi; Chytridiomycota; Neocallimastixales;

OC Neocallimastixaceae; *Pyromyces*.

OX NCBI_TaxID=45796;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-96313314; PubMed=8768520;

RA Milward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P., Gilbert H.J.;

RT "Evidence that the *Pyromyces* gene family encoding endo-1,4-mannanases arose through gene duplication."

RT FEMS Microbiol. Lett. 141:183-188(1996).

RL -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannan, galactomannans, glucomannans, and galactoglucomannans.

CC -1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANA, MANC AND XYNA; AND TO THOSE OF N.PATRICIARUM XYNA.

CC -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL; X97408; CAA6061.1; -

CC DR InterPro: IPR002883; CBD_5.

CC DR InterPro: IPR000805; Glyco_hydro_26.

CC DR Pfam: PF02013; CBD_5; 2.

CC DR Pfam: PF02156; Glyco_hydro_26; 1.

CC DR PRINTS: PR00739; GLYHYDRLASE26.

CC DR HYDROLASE; Glycosidase; Signal; Multigene family; Repeat.

CC KW SIGNAL 1 19 POTENTIAL.

CC FT CHAIN 20 571 MANMAN ENDO-1,4-BETA-MANNOSIDASE B.

CC FT DOMAIN 490 571 2 X 39 AA APPROXIMATE REPEATS.

CC FT REPEAT 490 527 1.

CC FT REPEAT 533 571 2.

CC FT DOMAIN 475 479 POLY-ASN.

CC FT DOMAIN 482 488 POLY-ASN.

CC FT SEQUENCE 571 AA; 64397 MW; B1344581FNA9DAA CRC64;

Query Match 75.0%; Score 42; DB 1; Length 571;

Best Local Similarity 60.0%; Pred. No. 4.3;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLSVPC 10
| | | | |
Db 492 CFSIPLGYPC 501

RESULT 3
IMH2_ARATH STANDARD; PRT: 502 AA.

AC Q9SA34;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable inosine 5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP dehydrogenase) (IMPDH) (IMPD).

GN ATG16350 OR F309.15.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.

OC NCBI_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RP STRAIN=CV. COLUMBIA;

RX MEDLINE-21016719; PubMed=11130712;

RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Malt R., Marshall A., Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Tatum M.J., Town C.D., Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;

RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."

RL Nature 408:816-820(2000).

CC -1- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O = xanthosine 5'-phosphate + NADH.

CC -1- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.

CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO GMP REDUCTASE.

CC -----

CC -1- SIMILARITY: CONTAINS 1 CBS DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL; AC006341; AAD34687.1; -

CC DR HSSP; P12268; 1830.

CC DR InterPro: IPR000644; CBS.

CC DR InterPro: IPR003009; FMN_enzyme.

CC DR InterPro: IPR001093; IMP_DH_GMP_RED.

CC DR Pfam; PF00571; CBS; 1.

CC DR Pfam; PF00478; IMPDH_C; 1.

CC DR Pfam; PF01574; IMPDH_N; 1.

CC DR PROSITE; PS00487; IMP_DH_GMP_RED; 1.

CC DR Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;

CC KW CBS domain. 167 219 CBS.

CC FT DOMAIN 167 219 CBS.

CC FT BINDING 321 321 IMP (POTENTIAL).

CC FT SEQUENCE 502 AA; 54051 MW; FB87D84160818310 CRC64;

Query Match 67.9%; Score 38; DB 1; Length 502;

Best Local Similarity 85.7%; Pred. No. 19;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 VPLSVPC 10
| | | | |
Db 49 VPLSVPC 55


```

RESULT 4
IMH1_ARATH STANDARD; PRT; 503 AA.
AC P47996;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205) (IMP
DE dehydrogenase) (IMPDH) (IMPD)
CN IMPDH OR A11G79470 OR T8K14.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN NM
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=97045815; PubMed=8890737;
RA Collart F.R., Osipuk J., Trent J., Olsen G.J., Huberman E.;
RT "Cloning and characterization of the gene encoding IMP dehydrogenase
RT from Arabidopsis thaliana."
RT Gene 174:217-220(1996).
[2]
RN RM
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altif H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chen Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson Hopson C., Khan S., Khaykin E.,
RA Kim C.U., Koo H.L., Kremetaskala I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltschker J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.U., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RT Nature 408:816-820(2000).
RL
CC -i- CATALYTIC ACTIVITY: Inosine 5'-phosphate + NAD(+) + H(2)O =
CC xanthosine 5'-phosphate + NADH.
CC -i- PATHWAY: FIRST REACTION UNIQUE TO GMP BIOSYNTHESIS.
CC -i- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO
CC GMP REDUCTASE.
CC -i- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L34684; AAB41940.1; -.
CC EMBL: AC007202; AAD30229.1; -.
CC HSSP: P12268; 1B30.
CC InterPro: IPR000644; CBS.
CC InterPro: IPR003009; FMN_enzyme.
CC InterPro: IPR001093; IMP_DH_GMP_RED.
CC Pfam: PF00571; CBS. 1.
CC Pfam: PF00478; IMPDH_C. 1.
CC Pfam: PF01574; IMPDH_N. 1.
CC PROSITE: PS00487; IMP_DH_GMP_RED; 1.

```

KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;
 KW CBS domain. 322 IMP (POTENTIAL);
 FT BINDING 322
 SQ SEQUENCE 503 AA; 54194 MW; ADDDAF9C3A697A9A CRC64;

Query Match 67.9%; Score 38; DB 1; Length 503;
 Best Local Similarity 85.7%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 5
Y396_HELPJ STANDARD; PRT; 616 AA.
ID Y396_HELPJ
AC Q92KF3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein JHP0985.
GN JHP0985.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN NM
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummuto P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Voyts G.F.,
RA "Genomic sequence comparison of two unrelated isolates of the human
RA gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -i- SIMILARITY: BELONGS TO THE UBIJ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001527; AAD06568.1; -.
CC DR InterPro: IPR002830; UPF0096.
CC DR Pfam: PF01977; UPF0096; 1.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 616 AA; 70989 MW; C219E1DCB4C8BD5 CRC64;

```

Query Match 67.9%; Score 38; DB 1; Length 616;
 Best Local Similarity 77.8%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 282 302 POTENTIAL.
 FT TRANSMEM 346 366 POTENTIAL.
 FT TRANSMEM 472 492 POTENTIAL.
 SQ SEQUENCE 581 AA; 64209 MW; B0592C460589319B CRC64;

Query Match 64.3%; Score 36; DB 1; Length 581;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLSPVC 10
 1:1:1:1:1
 DB 564 CVSIPVSSG 573

RESULT 9
 COTE_HUMAN
 ID CORE_HUMAN STANDARD; PRT; 669 AA.

AC P81408;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CORE1 protein.
 GN CORE1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HIPPOCAMPUS;
 RX MEDLINE=97474796; PubMed=9331372;
 RA Winfield S.L., Tavehi N., Martin B.M., Ginos E.I., Sidransky E.;
 RT "Identification of three additional genes contiguous to the
 RT glucocorticoidase locus on chromosome 12p1: implications for Gaucher
 RT disease.";
 RL Genome Res. 7:1020-1026(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: TO HUMAN KIA0574.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: AF023268; AAC51822.1; -
 KW Transmembrane.
 FT TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT DOMAIN 244 250 POLY-PRO.
 FT DOMAIN 635 639 POLY-SER.
 SQ SEQUENCE 669 AA; 71482 MW; E30360AC9A4571B6 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 669;
 Best Local Similarity 70.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISVPLSPVC 10
 1:1:1:1:1
 DB 139 CPSVPLLRPC 148

RESULT 10
 YMA2_MYCBO STANDARD; PRT; 298 AA.

AC 002278;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 32.3 kDa protein in MAS 5 region.
 OS Mycobacterium bovis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1765;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCG;
 RX MEDLINE=92406887; PubMed=1527058;
 RA Mathur M., Kolattukudy P.E.;
 RT "Molecular cloning and sequencing of the gene for mycosideric acid
 RT synthase, a novel fatty acid elongating multifunctional enzyme, from
 RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.";
 RL J. Biol. Chem. 267:19388-19395(1992).
 CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV1521.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: M95808; AAA25367.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 KW Hypothetical protein; Transmembrane.
 SQ SEQUENCE 298 AA; 32286 MW; 2232DD5DBC6F371 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 298;
 Best Local Similarity 87.5%; Pred. No. 38;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLSPVC 9
 1:1:1:1:1
 DB 63 IAVPLSPVC 70

RESULT 11
 ID MAF1_MOUSE STANDARD; PRT; 323 AA.

AC P54841;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor MAF1 (Segmentation protein KR) (Kreisler).
 GN MAFB OR MAF1 OR KRML.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95094266; PubMed=8001130;
 RA Cordes S.P., Barsh G.S.;
 RT "The mouse segmentation gene *kr* encodes a novel basic domain-leucine
 RT zipper transcription factor.";
 RL Cell 79:1025-1034(1994)

CC -1- FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDRAIN
 CC SEGMENTATION).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN.
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND
 CC IN THE CAUDAL HINDRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES),
 CC THE HIGH LEVEL DOMAIN EXHIBITS A SHARP ROSTRAL EDGE COINCIDENT
 CC WITH THE R4/R5 BOUNDARY AND A DIFFUSE CAUDAL EDGE LOCATED MIDWAY
 CC THROUGH R6.

```

CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L36435; AAA65689.1; -
DR HSSP; P05412; 1JUN.
DR TRANSFAC; T01439; -
DR MGD; MGI:104555; Mafp.
DR InterPro; IPR001871; bZIP.
DR SMART; SM00338; BRLZ; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 131 143 POLY-HIS.
FT DOMAIN 158 167 POLY-HIS.
FT DNA_BIND 238 264 BASIC MOTIF.
FT DOMAIN 266 287 LEUCINE-ZIPPER.
FT MUTAGEN 248 248 N->S: LOSS OF TRANSCRIPTIONAL ACTIVITY.
SQ SEQUENCE 323 AA; 35809 MW; D77AE07ABD9C2AD2 CRC64;

Query Match
Best Local Similarity 62.5%; Score 35; DB 1; Length 323;
Matches 6; Conservative 0; Pred. No. 41; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 57 STPLSTPC 64

RESULT 12
MAFI_RAT
ID MAFI_RAT STANDARD; PRT; 323 AA.
AC P54842;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor MAF1.
GN MAFB OR MAF1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=WTSTAR; TISSUE=liver;
RX MEDLINE=97190228; PubMed=9038383;
RA Sakai M., Imaki J., Yoshida K., Ogata A., Matsushima-Hibaya Y.,
RA Kuboki Y., Nishizawa M., Nishi S.;
RT "Rat maf related genes: specific expression in chondrocytes, lens and
RT spinal cord.";
RL Oncogene 14:745-750(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U56241; AAB50062.1; -
DR HSSP; P05412; 1JUN.
DR InterPro; IPR001871; bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.

```

```

KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 131 143 POLY-HIS.
FT DOMAIN 158 167 POLY-HIS.
FT DOMAIN 194 201 POLY-ALA.
FT DNA_BIND 238 264 BASIC MOTIF.
FT DOMAIN 266 287 LEUCINE-ZIPPER.
SQ SEQUENCE 323 AA; 35792 MW; 6E386340D1F840A5 CRC64;

Query Match
Best Local Similarity 62.5%; Score 35; DB 1; Length 323;
Matches 6; Conservative 0; Pred. No. 41; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SVPLSVPC 10
Db 57 STPLSTPC 64

RESULT 13
V181_FOWPV
ID V181_FOWPV STANDARD; PRT; 369 AA.
AC Q9J552;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Protein FPV181.
GN FPV181.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3813-3831(2000).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES A16 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF198100; AAF44525.1; -
DR InterPro; IPR004251; DUF230.
DR Pfam; PF03003; DUF230; 1.
SQ SEQUENCE 369 AA; 42081 MW; 380A71032C18BB99 CRC64;

Query Match
Best Local Similarity 62.5%; Score 35; DB 1; Length 369;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CISVPLSVPC 10
Db 164 CIGTPSLPC 173

RESULT 14
FD21_MYCTU
ID FD21_MYCTU STANDARD; PRT; 578 AA.
AC O50441;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative fatty-acid--CoA ligase fadd21 (EC 6.2.2.1.-) (Acy1-CoA
DE synthetase).
GN FADD21 OR RV1185C OR MT1222 OR MTV005.21C.
OS Mycobacterium tuberculosis.

```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornby S., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL010186; CAA15862.1; -
 CC EMBL: AE006999; AAK45479.1; -
 CC TIGR: MT1222;
 DR Tuberculist; RV1185c; -
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; FALSE_NEG.
 KW Hypothetical protein; Ligase; Fatty acid metabolism;
 KW Complete proteome.
 SO SEQUENCE 578 AA; 62756 MW; 698F8FE9E4BFA2F CRC64;

Query Match 62.5%; Score 35; DB 1; Length 578;
 Best Local Similarity 87.5%; Pred. No. 72;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLSVP 9
 I:|||||
 Db 88 IAVPLSVP 95

RESULT 15
 FD21_MYCLE
 ID FD21_MYCLE STANDARD; PRT; 579 AA.
 AC P54200; O9CC61;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative fatty-acid-CoA ligase fadD21 (EC 6.2.1.-) (Acyl-CoA
 DE synthetase).
 GN FAD21 OR MASC OR ML1234.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RA MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall P., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornby T., Jagels K., Lacroix C., McLean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO SEVERAL
 CC FRAMESHIFTS IN THE C-TERMINUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00010; AAA17075.1; ALT_FRAME.
 CC EMBL: AL583921; CAC31615.1; -
 CC Leproma; ML1234; -
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; FALSE_NEG.
 KW Hypothetical protein; Ligase; Fatty acid metabolism;
 KW Complete proteome.
 FT CONFLICT 179
 SO SEQUENCE 579 AA; 63102 MW; 16B8AB8384FBFA1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 579;
 Best Local Similarity 87.5%; Pred. No. 72;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLSVP 9
 I:|||||
 Db 88 IAVPLSVP 95

Search completed: June 17, 2002, 16:24:32
 Job time: 1561 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:44 ; Search time 172.85 Seconds
(without alignments)
10.008 Million cell updates/sec

Title: US-09-761-636a-13

Perfect score: 56
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvivirus:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	71.4	109	2	Q9L4Y0
2	39	69.6	85	2	Q9S556
3	39	69.6	85	16	Q9S564
4	39	69.6	148	16	Q9PBM2
5	39	69.6	620	12	Q89893
6	37	66.1	86	9	Q9MC88
7	37	66.1	89	16	Q9P9Y4
8	37	66.1	284	16	Q9ZRN5
9	37	66.1	316	5	Q9V7A3
10	37	66.1	355	10	Q9L6Y9
11	37	66.1	370	5	Q9VA44
12	37	66.1	501	10	Q9AY75
13	37	66.1	725	5	Q9N8V1
14	36	64.3	112	4	Q14634
15	36	64.3	134	11	Q63318
16	36	64.3	207	16	Q9CG60

17	36	64.3	208	5	Q9NF93
18	36	64.3	256	16	Q9RX44
19	36	64.3	280	10	Q94JK6
20	36	64.3	496	16	Q9HWV6
21	36	64.3	512	10	Q9FX29
22	36	64.3	668	4	Q9BR66
23	36	64.3	857	10	Q9ASK9
24	36	64.3	1136	13	Q98UH3
25	35	62.5	100	17	Q9Y998
26	35	62.5	144	4	Q9B214
27	35	62.5	161	5	Q9NB94
28	35	62.5	192	4	Q9B215
29	35	62.5	255	4	Q9B216
30	35	62.5	286	13	Q42290
31	35	62.5	286	13	Q57342
32	35	62.5	291	10	Q9PFV5
33	35	62.5	323	4	Q9Y5Q3
34	35	62.5	323	4	Q9H1F1
35	35	62.5	368	2	Q9K469
36	35	62.5	382	16	Q9HXJ9
37	35	62.5	438	10	P33452
38	35	62.5	503	16	Q9RV44
39	35	62.5	508	5	Q9N3R5
40	35	62.5	579	16	Q9CD79
41	35	62.5	580	16	P96290
42	35	62.5	582	2	P71495
43	35	62.5	706	16	Q83702
44	35	62.5	759	16	Q9Z955
45	35	62.5	1158	6	Q9TS23

ALIGNMENTS

RESULT 1
ID Q9L4Y0 PRELIMINARY; PRT; 109 AA.
AC Q9L4Y0;
DC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TCM F1 MONOOXYGENASE.
GN ELMH.
OS Streptomyces olivaceus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=47716;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TU 2353;
RX MEDLINE=21225632; PubMed=11325225;
RA Rafanan E.R. Jr., Le L., Zhao L., Decker H., Shen B.;
RT "Cloning, sequencing, and heterologous expression of the elmHJ genes
RT involved in the biosynthesis of the polyketide antibiotic elitoramycin
RT from Streptomyces olivaceus Tu2353.";
RL J. Nat. Prod. 64:444-449(2001).
DR EMBL; AF263463; AAF73051.1; -;
KW Monooxygenase.
SQ SEQUENCE 109 AA; 12304 MW; E65F4010CTD4B30B CRC64;

Query Match 71.4%; Score 40; DB 2; Length 109;
Best Local Similarity 77.8%; Pred. No. 4.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISVPLSVPC 10
Db 90 IATPLSVPC 98
RESULT 2
Q9S556 PRELIMINARY; PRT; 85 AA.

```

AC 09S556;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 9.3 KDA PROTEIN.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM14;
RX MEDLINE=20521925; PubMed=11069649;
RA Nakayama K., Takashima K., Ishihara H., Shinomiya T., Kageyama M.,
RA Kanaya S., Ohnishi M., Murata T., Mori H., Hayashi T.;
RT "The R-type pyocin of Pseudomonas aeruginosa is related to P2 phage,
RT and the F-type is related to lambda phage."
RL MOL. MICROBIOL. 38:213-231(2000).
DR EMBL; AB030826; BAA83139.1; -
KW Hypothetical protein.
SQ SEQUENCE 85 AA; 9252 MW; 472269231FCC6355 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 85;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 ISVPLSVC 10
Db 27 VEVPLAVPC 35

RESULT 3
ID 09S564 PRELIMINARY; PRT; 85 AA.
AC 09S564;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 9.2 KDA PROTEIN (HYPOTHETICAL PROTEIN PA0631).
GN PA0631.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RA Nakayama K., Takashima K., Ishihara H., Shinomiya T., Kageyama M.,
RA Kanaya S., Ohnishi M., Murata T., Terawaki Y., Mori H., Hayashi T.;
RT "Genetic relationship between bacteriophages and bacteriophages."
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SLM6;
RA Nakayama K., Takashima K., Ishihara H., Shinomiya T., Kageyama M.,
RA Kanaya S., Ohnishi M., Murata T., Mori H., Hayashi T.;
RT "The R-type pyocin is related to P2 phage, and the F-type pyocin is
RT related to lambda phage."
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltz L., Tolentino E., Westbrook-Watman S., Yan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).

```

```

DR EMBL; AB030825; BAA83170.1; -
DR EMBL; AB045308; BAA97449.1; -
DR EMBL; AE004498; AAG04020.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 85 AA; 9222 MW; F2267D9C542C7F48 CRC64;

Query Match 69.6%; Score 39; DB 16; Length 85;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 ISVPLSVC 10
Db 27 VEVPLAVPC 35

RESULT 4
ID 09PBM2 PRELIMINARY; PRT; 148 AA.
AC 09PBM2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF2118.
GN XF2118.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne W., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honelset J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pasquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL; AE004026; AAF84917.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 148 AA; 15990 MW; B78D2FAFB6B29260 CRC64;

Query Match 69.6%; Score 39; DB 16; Length 148;
Best Local Similarity 50.0%; Pred. No. 8.8;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLSVC 10
Db 8 CCAVPMNAPC 17

```


RESULT 5
ID 089893 PRELIMINARY; PRT; 620 AA.
AC 089893;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL 67.1 KDA PROTEIN.
GN DR2.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-95266321; PubMed-7747482;
RA Compels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efsthliou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution.";
RL Virology 209:29-51(1995).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-90080132; PubMed-2152817;
RA Lawrence G.L., Chee M., Craxton M.A., Compels U.A., Honess R.W.,
RA Barrett B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";
RL J. Virol. 64:287-299(1990).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-91237802; PubMed-1851860;
RA Chang C.K., Balachandran N.;
RT "Identification, characterization, and sequence analysis of a cDNA
RT encoding a phosphoprotein of human herpesvirus 6.";
RL J. Virol. 65:2884-2894(1991).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-91333007; PubMed-1651403;
RA Teo I.A., Griffin B.E., Jones M.D.;
RT "Characterization of the DNA polymerase gene of human herpesvirus 6.";
RL J. Virol. 65:4670-4680(1991).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-91226542; PubMed-1851252;
RA Thomson B.J., Efsthliou S., Honess R.W.;
RT "Acquisition of the human adeno-associated virus type-2 rep gene by
RT human herpesvirus type-6.";
RL Nature 351:78-80(1991).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-91374590; PubMed-1654446;
RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.;
RT "Identification of a transactivating function mapping to the putative
RT immediate-early locus of human herpesvirus 6.";
RL J. Virol. 65:5381-5390(1991).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-92333249; PubMed-1321206;
RA Efsthliou S., Lawrence G.L., Brown C.M., Barrell B.G.;
RT "Identification of homologues to the human cytomegalovirus US22 gene
RT family in human herpesvirus 6.";
RL J. Gen. Virol. 73:1661-1671(1992).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-92148942; PubMed-1310766;
RA Geng Y., Chandran B., Josephs S.F., Wood C.;
RT "Identification and characterization of a human herpesvirus 6 gene
RT segment that trans activates the human immunodeficiency virus type 1
RT promoter.";
RL J. Virol. 66:1564-1570(1992).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-93091236; PubMed-1333836;
RA Compels U.A., Carrs A.L., Sun N., Arind J.R.;
RT "Infectivity determinants encoded in a conserved gene block of human
RT herpesvirus-6.";
RL DNA Seq. 3:25-39(1992).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-1374813;
RA Neipel F., Ellinger K., Fleckenstein B.;
RT "Gene for the major antigenic structural protein (p100) of human
RT herpesvirus 6.";
RL J. Virol. 66:3918-3924(1992).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-92333248; PubMed-1321205;
RA Thomson B.J., Honess R.W.;
RT "The right end of the unique region of the genome of human herpesvirus
RT 6 U1102 contains a candidate immediate early gene enhancer and a
RT homologue of the human cytomegalovirus US22 gene family.";
RL J. Gen. Virol. 73:1649-1660(1992).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-93187613; PubMed-8383182;
RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Fiume G.,
RA Fleckenstein B.;
RT "The glycoprotein B homologue of human herpesvirus 6.";
RL J. Gen. Virol. 74:495-500(1993).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-93224882; PubMed-8385692;
RA Compels U.A., Carrigan D.R., Carrs A.L., Arno J.;
RT "Two groups of human herpesvirus 6 identified by sequence analyses of
RT laboratory strains and variants from Hodgkin's lymphoma and bone
RT marrow transplant patients.";
RL J. Gen. Virol. 74:613-622(1993).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-93389439; PubMed-8397282;
RA Liu D.X., Compels U.A., Nicholas J., Lelliott C.;
RT "Identification and expression of the human herpesvirus 6 glycoprotein
RT H and interaction with an accessory 40K glycoprotein.";
RL J. Gen. Virol. 74:1847-1857(1993).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-94025558; PubMed-7692666;
RA Liu D.X., Compels U.A., Foa-Tomasi L., Campadelli-Fiume G.;
RT "Human herpesvirus-6 glycoprotein H and L homologs are components of
RT the gp100 complex and the gp external domain is the target for
RT neutralizing monoclonal antibodies.";
RL Virology 197:12-22(1993).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE-93331710; PubMed-7687803;
RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,
RA Greenmayer C., Dambaugh T.R.;
RT "A strongly immunoreactive virion protein of human herpesvirus 6

RT variant B strain 229: Identification and characterization of the gene
RT and mapping of a variant-specific monoclonal antibody reactive
RT epitope.";
RN Virology 195:521-531(1993).
RN [17]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-93323202; PubMed-7687301;
RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwatnpong S.,
RT Chandran B.;
RT "Identification and mapping of the gene encoding the glycoprotein
RT complex gp82-gp105 of human herpesvirus 6 and mapping of the
RT neutralizing epitope recognized by monoclonal antibodies.";
RN J. Virol. 67:4611-4620(1993).
RN [18]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-95146989; PubMed-7844567;
RA Compels U.A., Macaulay H.A.;
RT "Characterization of human telomeric repeat sequences from human
RT herpesvirus 6 and relationship to replication.";
RN J. Gen. Virol. 76:451-458(1995).
RN [19]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-94047392; PubMed-8230490;
RA Demhurst S., Dollard S.C., Pellett P.E., Dambaugh T.R.;
RT "Identification of a lytic-phase origin of DNA replication in human
RT herpesvirus 6B strain 229.";
RN J. Virol. 67:7680-7683(1993).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-94118404; PubMed-8289364;
RA Nicholas J., Martin M.;
RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
RT genome of human herpesvirus 6 encoding human cytomegalovirus
RT immediate-early gene homologs and transactivating functions.";
RN J. Virol. 68:597-610(1994).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-94202284; PubMed-8151768;
RA Schiwe U., Neipel F., Schreiner D., Fleckenstein B.;
RT "Structure and transcription of an immediate-early region in the human
RT herpesvirus 6 genome.";
RN J. Virol. 68:2978-2985(1994).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN-01102;
RX MEDLINE-94181269; PubMed-8134119;
RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
RA Frenkel N., Rosenthal L.J.;

Query Match 69.6%; Score 39; DB 12; Length 620;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CISVPLSVP 9
Db 439 CPSIPLSVP 447

RESULT 5
09MC88 PRELIMINARY; PRT; 86 AA.
AC 09MC88;
DT 01-OCT-2000 (TRENBLREL. 15, Created)

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE ORE33.
GN ORE33.
OS Bacteriophage D3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=31535;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20042341; PubMed-10572124;
RA Gilakjan Z.A., Kropinski A.M.;
RT "Cloning and analysis of the capsid morphogenesis genes of Pseudomonas
RT aeruginosa bacteriophage D3: another example of protein chain malty?";
RN J. Bacteriol. 181:7221-7227(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20485557; PubMed-11029426;
RA Kropinski A.M.;
RT "Sequence of the Genome of the Temperate, Serotype-convertible,
RT Pseudomonas aeruginosa Bacteriophage D3.";
RN J. Bacteriol. 182:6066-6074(2000).
DR EMBL: AF165214; AAF80793.1; -;
SQ SEQUENCE 86 AA; 9389 MW; 86ABCA82B2A3E59F CRC64;

Query Match 65.1%; Score 37; DB 9; Length 86;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 ISVPLSVP 10
Db 28 VEVPIAVPC 36

RESULT 7
09P9Y4 PRELIMINARY; PRT; 89 AA.
AC 09P9Y4;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XP2736.
GN XP2736.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9A5C;
RX MEDLINE-20365717; PubMed-10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Medeira A.M.B.N., Medeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nnaní A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Valla da H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Weidenis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-159(2000).
 DR EMBL; AE004080; AAF85521.1; -;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 89 AA; 9641 MW; F13175F1054D6130 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 89;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 ISVPLSVP 10
 DB 19 VSAFLRVP 27

RESULT 8
 ID 092RNS PRELIMINARY; PRT; 284 AA.
 AC 092RNS:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
 GN SMC00879.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_Taxid=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE-21368234; PubMed-11474104;
 RA Gallibert F., Flanin T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Baurly-Hubler F., Barrett M.J., Becker A., Boistard P., Bothé G.,
 RA Baurly M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chalm P.,
 RA Gloux S., Godt T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masny D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 RA Ramsperger U., Surzycki R., Thebaud P., Vandenbol M.,
 RA Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont *Sinorhizobium meliloti*.";
 RL Science 293:668-672(2001).
 DR EMBL; AL591785; CAC45399.1; -;
 KM Complete proteome.
 SQ SEQUENCE 284 AA; 31448 MW; BC33BF3FE0AE6AB2 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 284;
 Best Local Similarity 55.6%; Pred. No. 38;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 CISVPLSVP 9
 DB 211 CVAAPISVP 219

RESULT 9
 ID 09V743 PRELIMINARY; PRT; 316 AA.
 AC 09V743:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE CG12860 PROTEIN.
 GN CG12860.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
 OC Pelegruta; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan D.A., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003813; AAF58222.1; -;
 DR FlyBase; FBgn0033954; CG12860.
 SQ SEQUENCE 316 AA; 36644 MW; B2703C1D4C21969 CRC64;

Query Match 66.1%; Score 37; DB 5; Length 316;
 Best Local Similarity 50.0%; Pred. No. 42;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CISVPLSVP 10
 DB 205 CVSAPKAPC 214

RESULT 10
 ID 09LGY9 PRELIMINARY; PRT; 355 AA.
 AC 09LGY9:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE HYPOTHETICAL PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.

NCBI_TaxID=4530;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV_NIPONBARE;
 RT Sasaki T., Matsunoto T., Yamamoto K.;
 "Oryza sativa niponbare(Ga3) genomic DNA, chromosome 1, PAC
 clone: p0702f03."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002481; BAA96584.1; -
 KW Hypothetical protein.
 SO SEQUENCE 355 AA; 38405 MW; DB831543779C815E CRC64;

Query Match 66.1%; Score 37; DB 10; Length 355;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CISPVSPLSPC 10
 : : : : :
 Db 48 CVCGLPLPPC 57

RESULT 11
 Q9VA44 PRELIMINARY; PRT; 370 AA.
 AC Q9VA44;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CG11313 PROTEIN.
 GN CG11313.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkov D., Botchan M.R., Bock J., Brokstein P., Brotler P.,
 Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de la Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson R.K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Hariri N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei V., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier-Berger, Spadling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AE003775; AAF57079.1; -
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.09A.1.
 DR Flybase: FBgn0039798; CG11313.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYP_SEC.1.
 DR PROSITE: PS50240; TRYPsin_DOM.1.
 DR PROSITE: PS00134; TRYPsin_HIS.1.
 DR PROSITE: PS00135; TRYPsin_SER.1.
 DR Hydrolase: Serine protease.
 SO SEQUENCE 370 AA; 40937 MW; E1E54E03F2CDE2AA CRC64;

Query Match 66.1%; Score 37; DB 5; Length 370;
 Best Local Similarity 55.6%; Pred. No. 49;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISPVSPLSP 9
 : : : : :
 Db 34 CVCNIPLCVP 42

RESULT 12
 Q9AY75 PRELIMINARY; PRT; 501 AA.
 AC Q9AY75;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE INOSINE MONOPHOSPHATE DEHYDROGENASE.
 GN OSJNA009119.1
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; Oryza.
 NC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_NIPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hall J.N., Burr P.C., Hsiao J.,
 Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Uteberg T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNA009119 genomic sequence."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC084320; AAR09225.1; -
 DR HSSP: P12268; IB30.
 DR InterPro: IPR000644; CBS.
 DR InterPro: IPR003009; FMN enzyme.
 DR InterPro: IPR001093; IMP_DH_GMP_RED.
 DR Pfam: PF00478; IMPDH_C.1.
 DR Pfam: PF01574; IMPDH_N.1.
 DR SMART: SM00116; CBS.1.
 DR PROSITE: PS00487; IMP_DH_GMP_RED.1.
 SO SEQUENCE 501 AA; 52655 MW; DDE295CB714C3802 CRC64;

Query Match 66.1%; Score 37; DB 10; Length 501;
 Best Local Similarity 71.4%; Pred. No. 65;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPLSVP 10
 : : : : :
 Db 4 VPLSVP 10

Db 54 IPISTPC 60

RESULT 13

ID 09NBV1 PRELIMINARY; PRT: 725 AA.

AC 09NBV1; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE POSSIBLE N-ETHYLALANINE SENSITIVE FUSION PROTEIN.
GN CHR1.136.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRE0927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Leonard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: A1359782; CAB5432.1; -
DR HSP: P18708; INSE.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003960; AAA_sub.
DR InterPro: IPR003959; AAA_subfam.
DR Pfam: PF00004; AAA_2.
DR SMART: SM00382; AAA_1.
DR PROSITE: PS00674; AAA_1.
KW ATP-binding.
SQ SEQUENCE 725 AA; 79954 MW; 37F1F12446260498 CRC64;

Query Match

Best Local Similarity 66.1%; Score 37; DB 5; Length 725;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISEPLSVPC 10

Db 642 CFSVKMHPG 651

RESULT 14

ID 014634 PRELIMINARY; PRT: 112 AA.

AC 014634; 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SKIN-SPECIFIC PROTEIN (FRAGMENT).
GN XP32.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RX MEDLINE=98008911; PubMed=9344646;
RA Zhao X.P., Elder J.T.;
RT "Positional cloning of novel skin-specific genes from the human
RT epidermal differentiation complex.";
RL Genomics 45:250-258(1997).
DR EMBL: AF005081; AAB83961.1; -
DR InterPro: IPR000561; EGF-like.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 112 AA; 11329 MW; 96826DC3E7362E6B CRC64;

Query Match 64.3%; Score 36; DB 4; Length 112;
Best Local Similarity 50.0%; Pred. No. 24;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISEPLSVPC 10
Db 27 CVTDCSAPC 36

RESULT 15

ID 063318 PRELIMINARY; PRT: 134 AA.

AC 063318; 063319; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RAT LY6-C ANTIGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90152758; PubMed=2154400;
RA Friedman S., Palfree R.G.E., Strlin S., Haemmerling U.;
RT "Analysis of three distinct Ly6-A-related cDNA sequences isolated from
RT rat kidney."
RL Immunogenetics 31:104-111(1990).
DR EMBL: M30690; AAA41547.1; -
DR EMBL: M30691; AAA41548.1; -
DR HSP: P01468; ICYN.
DR InterPro: IPR003632; Ly-6_CD59.
DR InterPro: IPR001526; Ly6_UPAR.
DR Pfam: PF00021; UPAR_LY6_1.
DR ProDom: PD003128; Ly-6_CD59; 1.
DR SMART: SM00134; LU; 1.
SQ SEQUENCE 134 AA; 14157 MW; 58D6ACDB457A7B9 CRC64;

Query Match

Best Local Similarity 64.3%; Score 36; DB 11; Length 134;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 CISEPLSVPC 10

Db 32 CIEVPLNANC 41

Search completed: June 17, 2002, 16:23:45
Job time: 1619 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:16 ; Search time 227.35 Seconds
(without alignments)
4.886 Million cell updates/sec

Title: US-09-761-636a-13
Perfect score: 56
Sequence: 1 C1SVPLSVPC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	22	AAU04532
2	45.5	81.2	11	22	AAU04526
3	45	80.4	141	16	AAU04534
4	43.5	77.7	11	22	AAU04542
5	41.5	74.1	9	22	AAU04533
6	41.5	74.1	11	22	AAU04545
7	40.5	72.3	11	22	AAU04544
8	39	69.6	77	22	AAU05534
9	39	69.6	108	22	AAU05827
10	38.5	68.8	11	22	AAU04543
11	38	67.9	59	22	ABG12530

12	38	67.9	59	22	ABG12542	Novel human diagno
13	38	67.9	71	22	ABM20407	Protein #2406 enco
14	38	67.9	71	22	AAU16003	Peptide #437 enco
15	38	67.9	502	21	AAU30888	Arabidopsis thalia
16	38	67.9	503	21	AAU20989	Arabidopsis thalia
17	38	67.9	503	21	AAU43108	Arabidopsis thalia
18	37	66.1	108	22	ABU15040	Human nervous syst
19	37	66.1	316	22	ABU67498	Drosophila melanog
20	37	66.1	368	22	ABG21894	Novel human diagno
21	37	66.1	370	22	ABU65635	Novel human diagno
22	37	66.1	438	22	ABG20151	Novel human diagno
23	36	64.3	9	17	AAU96138	Protease substrate
24	36	64.3	9	19	AAU82212	D-NorPEPS-A proteas
25	36	64.3	9	19	AAU45652	Peptide backbone o
26	36	64.3	63	22	AAU58478	Protonibacterium
27	36	64.3	65	21	AAU45459	Human secreted pro
28	36	64.3	65	21	AAU45460	Human secreted pro
29	36	64.3	76	22	AAU89148	Human secreted pro
30	36	64.3	83	21	AAU45461	Human secreted pro
31	36	64.3	85	21	AAU45462	Human secreted pro
32	36	64.3	109	22	ABU19452	Novel human diagno
33	36	64.3	111	22	AAU6053	Protonibacterium
34	36	64.3	121	21	AAU04076	Human secreted pro
35	36	64.3	125	22	AAU79027	Human protein SEQ
36	36	64.3	148	22	AAU80011	Human protein SEQ
37	36	64.3	208	22	ABG11773	Novel human diagno
38	35.5	63.4	80	21	AAU87243	Human signal pepti
39	35.5	63.4	80	22	AAU29311	Human PRO polypept
40	35	62.5	28	21	AAU44963	Human secreted pro
41	35	62.5	42	21	AAU44369	Human secreted pro
42	35	62.5	43	22	ABU35159	Peptide #265 enco
43	35	62.5	43	22	ABU20590	Protein #2589 enco
44	35	62.5	43	22	AAU55992	Human brain expres
45	35	62.5	43	22	AAU65360	Human bone marrow

ALIGNMENTS

RESULT 1	AAU04532	standard; Peptide: 10 AA.
ID	AAU04532	
AC	AAU04532	
XX		
XX	26-SEP-2001	(first entry)
XX		
DE	VEGF based monocyclic peptide 10.	
XX		
KW	Human; VEGF; vascular endothelial growth factor; angiogenesis;	
KW	neovascularisation; lymphangiogenesis; psoriasis; tumour;	
KW	diabetes induced neovascular sequelae; rheumatoid arthritis;	
KW	diabetic retinopathy; chronic inflammation; cyclic.	
XX		
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	Disulfide bond	1..10
FT	/note="This bond cyclises the peptide"	
PN	WO200152875-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	18-JAN-2001; 2001WO-US01533.	
XX		
PR	18-JAN-2000; 2000US-0176293.	
XX		
PR	16-MAY-2000; 2000US-0204590.	
XX		
PA	(LUDW-) LUDWIG INST CANCER RES.	
XX		
XX	Achen MG, Hughes RA, Stackner S, Cendron A;	
PI		
XX		

DR WPI: 2001-442248/47.
XX Novel monomeric monocyclic peptide, used to interfere with
XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
XX Claim 49; Page 32; 102pp; English.
XX
XX The sequence represents a monomeric monocyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
CC VEGFD (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
XX Sequence 10 AA:
SQ
Query Match 100.0%; Score 56; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 C1SVPLSVPC 10
Db 1 C1SVPLSVPC 10
RESULT 2
AAU04526
ID AAU04526 standard; Peptide: 11 AA.
XX
XX AAU04526;
XX
XX 26-SEP-2001 (first entry)
XX
XX VEGF based monocyclic peptide 3.
XX
XX Human; VEGF; vascular endothelial growth factor; angiogenesis;
XX neovascularisation; lymphangiogenesis; psoriasis; tumour;
XX diabetes induced neovascular sequelae; rheumatoid arthritis;
XX diabetic retinopathy; chronic inflammation; cyclic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Disulfide-bond 1..11
XX FT /note="This bond cyclises the peptide"
XX
XX W0200152875-A1.
XX
XX 26-JUL-2001.
XX
XX

XX
XX 18-JAN-2001; 2001WO-US01533.
XX
XX
XX 18-JAN-2000; 2000US-0176293.
XX
XX 16-MAY-2000; 2000US-0204590.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Hughes RA, Staeker S, Cendron A;
XX
XX WPI: 2001-442248/47.
XX
XX Novel monomeric monocyclic peptide, used to interfere with
XX angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
XX Claim 49; Page 32; 102pp; English.
XX
XX The sequence represents a monomeric monocyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
CC VEGFD (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
XX Sequence 11 AA:
SQ
Query Match 81.2%; Score 45.5; DB 22; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 C1SVPL-SVPC 10
Db 1 C1SVPLSVPC 11
RESULT 3
AAR66394
ID AAR66394 standard; Peptide: 141 AA.
XX
XX AAR66394;
XX
XX 11-AUG-1995 (first entry)
XX
XX Human psychosis protecting peptide encoded by reading frame 1.
XX
XX Psychosis protecting peptide.
XX
XX Homo sapiens.
XX
XX

XX	FT	Key	Location/Qualifiers	XX
XX	FT	Misc-difference 1..141	/note- "X corresp. to translated stop codon"	XX
XX	PN	W09426107-A.		XX
XX	PD	24-NOV-1994.		XX
XX	PF	13-MAY-1994;	94WO-US05445.	XX
XX	PR	13-MAY-1993;	93US-0060560.	XX
XX	PA	(UYNY) UNIV NEW YORK STATE.		XX
XX	PI	Basham DA, Friedhoff AJ, Miller JC;		XX
XX	DR	WPI: 1995-006234/01.		XX
XX	DR	N-PSDB; AA075090.		XX
XX	PT	New nucleic acids encoding psychosis protecting peptide and		XX
XX	PT	antibodies - for the treatment, diagnosis and research of		XX
XX	PT	psychotic disorders, such as schizophrenia		XX
XX	PS	Claim 1; Page 56; 87 pp: English.		XX
XX	CC	Psychosis protecting peptide (PP peptide) can be obtd. Initially		XX
XX	CC	by using the sequence in AA075090 as a basis for designing		XX
XX	CC	polynucleotide probes to clone, sequence and express or synthesize		XX
XX	CC	PP related proteins and peptides occurring in normal individuals,		XX
XX	CC	and to a substantially lesser degree in individuals with		XX
XX	CC	psychotic disorders. The PP encoding gene was discovered using		XX
XX	CC	the subcloning of cDNA from mRNA obtd. from monozygotic		XX
XX	CC	twins discordant for schizophrenia and assaying for clones		XX
XX	CC	in which expression is greatest in the 'well' twin. A nt		XX
XX	CC	sequence substantially corresp. to AA075090 is claimed, which		XX
XX	CC	comprises 30-x nts, where x=80, 95, 158, 222, 243, 249, 260, 295,		XX
XX	CC	407 or 423. A PP peptide of 10-141 AAs is also claimed.		XX
XX	SO	Sequence 141 AA;		XX
XX	QY	Query Match	80.4%; Score 45; DB 16; Length 141;	XX
XX	QY	Best Local Similarity	77.8%; Pred. No. 4.4;	XX
XX	QY	Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;		XX
XX	QY	1 C1SVPLSVP 9		XX
XX	QY	: :		XX
XX	Db	28 c1splsip 36		XX
XX	RESULT 4			XX
XX	AAU04542			XX
XX	ID	AAU04542 standard; Peptide: 11 AA.		XX
XX	AC	AAU04542;		XX
XX	DT	26-SEP-2001 (first entry)		XX
XX	DE	VEGF based monocyclic peptide 20.		XX
XX	KW	Human; VEGF: vascular endothelial growth factor; angiogenesis;		XX
XX	KW	neovascularisation; lymphangiogenesis; psoriasis; tumour;		XX
XX	KW	diabetes induced neovascular sequelae; rheumatoid arthritis;		XX
XX	KW	diabetic retinopathy; chronic inflammation; cyclic.		XX
XX	OS	Synthetic.		XX
XX	FT	Key	Location/Qualifiers	XX
XX	FT	Disulfide-bond 1..11		XX
XX	FT	/note- "This bond cyclises the peptide"		XX
XX	PN	W0200152875-A1.		XX

XX		26-JUL-2001.
XX	PE	18-JAN-2001; 2001WO-US01533.
XX	PR	18-JAN-2000; 2000US-0176293.
XX	PR	16-MAY-2000; 2000US-0204590.
XX	PA	(LUDW-) LUDWIG INST CANCER RES.
PI	Achen MG,	Hughes RA, Stacker S, Cendon A;
DR	WPI:	2001-442248/47.
XX		
PT		Novel monomeric monocyclic peptide, used to interfere with
PT		angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT		loop fragment from an exposed loop of a growth factor protein by
PT		oxidizing the cysteine residues -
XX		
PS	Example 25;	Page 47; 102pp; English.
CC		The sequence represents a monomeric monocyclic peptide of the invention,
CC		whose 3-dimensional structure is modelled on the expose loop of human
CC		VSEFD (vascular endothelial growth factor). The invention relates to a
CC		method of producing a monomeric monocyclic peptide by a measuring
CC		beta-beta carbon separation distances on opposite antiparallel strands of
CC		a peptide loop fragment from an exposed loop of a growth factor protein
CC		and cyclising the peptide by oxidising the cysteine residues. The
CC		monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC		monocyclic peptides) and a cyclic peptide with at least one amino acid
CC		deleted prior to cyclisation are used to interfere with angiogenesis,
CC		neovascularisation or lymphangiogenesis in a mammal with a condition
CC		characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC		The condition is diabetic retinopathy, psoriasis, arthropathy,
CC		hemangioma, vascularised malignant or benign tumour, post-recovery
CC		cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC		trauma, substance-induced neovascularisation of the liver, excessive
CC		hormone-related angiogenic dysfunction, diabetes induced neovascular
CC		sequelae, hypertension induced neovascular sequelae, or chronic liver
CC		infection. The peptides are also used to modulate vascular permeability
CC		in a mammal (the mammal has a condition characterised by fluid
CC		accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC		or brain. The peptides are used to image blood vessels and lymphatic
CC		vasculature. The monomeric and bicyclic peptides are used to interfere
CC		with at least one biological activity induced by VEGF, bFGF-C or -D and
CC		are also used in combination with an anti-inflammatory agent, to treat a
CC		chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC		diabetic retinopathy.
SQ	Sequence	11 AA;
QY	Query Match	77.7%, Score 43.5; DB 22; Length 11;
Bst	Local Similarity	81.8%; Pred. No. 0.55;
Matches	9; Conservative	1; Mismatches 0; Indels 1; Gaps 1;
Db	1 CISVPL-SVPC 10	
	:-	
	1 CISVPLSVPC 11	
RESULT	5	
ID	AAU04533	
AAU04533	standard; Peptide: 9 AA.	
AC	AAU04533;	
DT	26-SEP-2001 (first entry)	
DE	VEGF based monocyclic peptide 11.	
KM	Human: VEGF; vascular endothelial growth factor; angiogenesis;	
KM	neovascularisation; lymphangiogenesis; psoriasis; tumour;	
KM	diabetes induced neovascular sequelae; rheumatoid arthritis;	

KW diabetic retinopathy; chronic inflammation; cyclic.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..9 /note="This bond cyclises the peptide"
XX
XX MO200152875-A1.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01533.
XX
XX 18-JAN-2000; 2000US-0176293.
XX 16-MAY-2000; 2000US-0204590.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Hughes RA, Stacker S, Cendron A;
XX
XX WPI: 2001-442248/47.
XX
XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
XX
XX Claim 49; Page 32; 102pp; English.
XX
XX The sequence represents a monomeric monocyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
XX VEGFD (vascular endothelial growth factor). The invention relates to a
XX method of producing a monomeric monocyclic peptide by a measuring
XX beta-beta carbon separation distances on opposite antiparallel strands of
XX a peptide loop fragment from an exposed loop of a growth factor protein
XX and cyclising the peptide by oxidising the cysteine residues. The
XX monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX monocyclic peptides) and a cyclic peptide with at least one amino acid
XX deleted prior to cyclisation are used to interfere with angiogenesis,
XX neovascularisation or lymphangiogenesis in a mammal with a condition
XX characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX The condition is diabetic retinopathy, psoriasis, arthropathy,
XX hemangioma, vascularised malignant or benign tumour, post-recovery
XX cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX trauma, substance-induced neovascularisation of the liver, excessive
XX hormone-related angiogenic dysfunction, diabetes induced neovascular
XX sequelae, hypertension induced neovascular sequelae, or chronic liver
XX infection. The peptides are also used to modulate vascular permeability
XX in a mammal (the mammal has a condition characterised by fluid
XX accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX or brain. The peptides are used to image blood vessels and lymphatic
XX vasculature. The monomeric and bicyclic peptides are used to interfere
XX with at least one biological activity induced by VEGF, VEGF-C or -D and
XX are also used in combination with an anti-inflammatory agent, to treat a
XX chronic inflammation, especially rheumatoid arthritis, psoriasis and
XX diabetic retinopathy.
XX
XX Sequence 9 AA;
XX
XX
XX Query Match 74.1%; Score 41.5; DB 22; Length 9;
XX Best Local Similarity 90.0%; Pred. No. 6.4e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX 1 C1SVPLSYVC 10
XX ||||| |||
XX 1 C1SVPL-VPC 9
XX
XX
XX RESULT 6
XX AAU04545
XX ID AAU04545 standard; Peptide: 11 AA.

XX
XX AAU04545;
XX
XX 26-SEP-2001 (first entry)
XX
XX VEGF based monocyclic peptide 23.
XX
XX Human; VEGF; vascular endothelial growth factor; angiogenesis;
XX neovascularisation; lymphangiogenesis; psoriasis; tumour;
XX diabetes induced neovascular sequelae; rheumatoid arthritis;
XX diabetic retinopathy; chronic inflammation; cyclic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1..11 /note="This bond cyclises the peptide"
XX
XX
XX MO200152875-A1.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01533.
XX
XX 18-JAN-2000; 2000US-0176293.
XX 16-MAY-2000; 2000US-0204590.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Hughes RA, Stacker S, Cendron A;
XX
XX WPI: 2001-442248/47.
XX
XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
XX
XX Example 25; Page 47; 102pp; English.
XX
XX The sequence represents a monomeric monocyclic peptide of the invention,
XX whose 3-dimensional structure is modelled on the expose loop of human
XX VEGFD (vascular endothelial growth factor). The invention relates to a
XX method of producing a monomeric monocyclic peptide by a measuring
XX beta-beta carbon separation distances on opposite antiparallel strands of
XX a peptide loop fragment from an exposed loop of a growth factor protein
XX and cyclising the peptide by oxidising the cysteine residues. The
XX monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX monocyclic peptides) and a cyclic peptide with at least one amino acid
XX deleted prior to cyclisation are used to interfere with angiogenesis,
XX neovascularisation or lymphangiogenesis in a mammal with a condition
XX characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX The condition is diabetic retinopathy, psoriasis, arthropathy,
XX hemangioma, vascularised malignant or benign tumour, post-recovery
XX cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX trauma, substance-induced neovascularisation of the liver, excessive
XX hormone-related angiogenic dysfunction, diabetes induced neovascular
XX sequelae, hypertension induced neovascular sequelae, or chronic liver
XX infection. The peptides are also used to modulate vascular permeability
XX in a mammal (the mammal has a condition characterised by fluid
XX accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX or brain. The peptides are used to image blood vessels and lymphatic
XX vasculature. The monomeric and bicyclic peptides are used to interfere
XX with at least one biological activity induced by VEGF, VEGF-C or -D and
XX are also used in combination with an anti-inflammatory agent, to treat a
XX chronic inflammation, especially rheumatoid arthritis, psoriasis and
XX diabetic retinopathy.
XX
XX Sequence 11 AA;
XX
XX
XX Query Match 74.1%; Score 41.5; DB 22; Length 11;
XX Best Local Similarity 72.7%; Pred. No. 1.2;

Matches 8; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
OY 1 CISVPL-SVPC 10
|:||||:||||
Db 1 cvsylpttvc 11

RESULT 7
AA004544
ID AA004544 standard; Peptide; 11 AA.
XX
AC AA004544;
XX
DT 26-SEP-2001 (first entry)
XX
DE VEGF based monocyclic peptide 22.
XX
KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW diabetes induced neovascular sequelae; rheumatoid arthritis;
KW diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..11
FT /note="This bond cyclises the peptide"
XX
PN WO200152875-A1.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01533.
XX
PR 18-JAN-2000; 2000US-0176293.
PR 16-MAY-2000; 2000US-0204590.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Hughes RA, Stacker S, Cendron A;
XX
DR WPI; 2001-442248/47.
XX
PT Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
PS Example 25; Page 47; 102pp; English.
XX
CC The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring
CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere

CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
SO Sequence 11 AA;

Query Match 72.3%; Score 40.5; DB 22; Length 11;
Best Local Similarity 72.7%; Pred. No. 1.7;
Matches 8; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

OY 1 CISVPL-SVPC 10
|:||||:||||
Db 1 cisplissvpc 11

RESULT 8
AAM95534
ID AAM95534 standard; Protein; 77 AA.
XX
AC AAM95534;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 4192.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
XX Homo sapiens.
XX
PN WO20015320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214686.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 22-AUG-2000; 2000US-0226683.
PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251866.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259676.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI: 2001-465570/50.
 N-PSDB: AAL01504.
 Isolated nucleic acid molecule encoding a reproductive system antigen
 is used in preventing, treating or ameliorating a medical condition -
 PS Claim 11; SEQ ID NO 4192; 1297bp + Sequence Listing; English.
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 SQ Sequence 77 AA.
 Query Match 69.6%; Score 39; DB 22; Length 77;
 Best Local Similarity 40.0%; Pred. No. 22;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CTSVPLSYPC 10
 I::I::I
 Db 47 cInapiripic 56
 RESULT 9
 ID AAM95827 standard; Protein; 108 AA.
 AC AAM95827;
 XX
 XX 21-NOV-2001 (first entry)
 DE Human reproductive system related antigen SEQ ID NO: 4485.
 DE Human reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX

PN WO200155320-A2.
 XX 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US01339.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227099.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM.
 PI

XX WPI: 2001-465570/50.
DR N-PSDB: AAL01797.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 11; SEQ ID NO 4485; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a protein of the invention.
XX
SQ Sequence 108 AA;

Query Match 69.6%; Score 39; DB 22; Length 108;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISVPLSVPC 10
Db 33 vsvplshpc 41

RESULT 10
AAU04543
ID AAU04543 standard; Peptide: 11 AA.
XX
AC AAU04543;
XX
DT 26-SEP-2001 (first entry)
XX
DE VEGF based monocyclic peptide 21.
XX
XX Human: VEGF: vascular endothelial growth factor; angiogenesis;
KM neovascularisation; lymphangiogenesis; psoriasis; tumour;
KM diabetes induced neovascular sequelae; rheumatoid arthritis;
KM diabetic retinopathy; chronic inflammation; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..11
FT /note- "This bond cyclises the peptide"
XX
XX WO200152875-A1.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01533.
XX
XX 18-JAN-2000; 2000US-0176293.
XX
XX 16-MAY-2000; 2000US-0204590.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Achen MG, Hughes RA, Stacker S, Cendron A;
XX
XX WPI: 2001-442248/47.
XX
XX Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX
XX Example 25; Page 47; 102pp; English.
XX
CC The sequence represents a monomeric monocyclic peptide of the invention,
CC whose 3-dimensional structure is modelled on the expose loop of human
CC VEGF (vascular endothelial growth factor). The invention relates to a
CC method of producing a monomeric monocyclic peptide by a measuring

CC beta-beta carbon separation distances on opposite antiparallel strands of
CC a peptide loop fragment from an exposed loop of a growth factor protein
CC and cyclising the peptide by oxidising the cysteine residues. The
CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
CC monocyclic peptides) and a cyclic peptide with at least one amino acid
CC deleted prior to cyclisation are used to interfere with angiogenesis,
CC neovascularisation or lymphangiogenesis in a mammal with a condition
CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
CC The condition is diabetic retinopathy, psoriasis, arthropathy,
CC hemangioma, vascularised malignant or benign tumour, post-recovery
CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
CC trauma, substance-induced neovascularisation of the liver, excessive
CC hormone-related angiogenic dysfunction, diabetes induced neovascular
CC sequelae, hypertension induced neovascular sequelae, or chronic liver
CC infection. The peptides are also used to modulate vascular permeability
CC in a mammal (the mammal has a condition characterised by fluid
CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
CC or brain. The peptides are used to image blood vessels and lymphatic
CC vasculature. The monomeric and bicyclic peptides are used to interfere
CC with at least one biological activity induced by VEGF, VEGF-C or -D and
CC are also used in combination with an anti-inflammatory agent, to treat a
CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
CC diabetic retinopathy.
XX
SQ Sequence 11 AA;

Query Match 68.8%; Score 38.5; DB 22; Length 11;
Best Local Similarity 63.6%; Pred. No. 3.7;
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 CISVPL-SVPC 10
Db 1 citplshpc 11

RESULT 11
ABG12530
ID ABG12530 standard; Protein: 59 AA.
XX
AC ABG12530;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #12521.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX N-PSDB: AAS76717.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 42889; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 59 AA;

Query Match 67.9%; Score 38; DB 22; Length 59;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 C1SVPLSVPC 10
||||: |||
Db 9 C1SVPLSVPC 18

RESULT 12
ABG12542
ID ABG12542 standard; Protein; 59 AA.
XX
AC ABG12542;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12533.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS76729.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 42901; 103pp; English.
XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 59 AA;

Query Match 67.9%; Score 38; DB 22; Length 59;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 C1SVPLSVPC 10
||||: |||
Db 9 C1SVPLSVPC 18

RESULT 13
ABB20407
ID ABB20407 standard; Protein; 71 AA.
XX
AC ABB20407;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #2406 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX

PS Claim 15; SEQ ID No 22177; 530bp; English.

XX The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC AB21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SO Sequence 71 AA;

Query Match 67.9%; Score 38; DB 22; Length 71;

Best Local Similarity 77.8%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CISVPLSVP 9

Db 51 cisvplptp 59

RESULT 14

AAM16003

ID AAM16003 standard; Protein; 71 AA.

XX

AC AAM16003;

XX

DT 12-OCT-2001 (first entry)

DE Peptide #2437 encoded by probe for measuring cervical gene expression.

XX

KW Probe: human; microarray; gene expression; cervical epithelial cell;

KM cervical cancer.

XX

OS Homo sapiens.

OS

PN WO200157278-A2.

PN

XX

PD 09-AUG-2001.

PD

XX

PF 30-JAN-2001; 2001WO-US00670.

PF

XX

PR 04-FEB-2000; 2000US-0180312.

PR

XX

PR 26-MAY-2000; 2000US-0207456.

PR

XX

PR 30-JUN-2000; 2000US-0608408.

PR

XX

PR 03-AUG-2000; 2000US-0632366.

PR

XX

PR 21-SEP-2000; 2000US-0234687.

PR

XX

PR 27-SEP-2000; 2000US-0236359.

PR

XX

PR 04-OCT-2000; 2000GB-0024263.

PR

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

PA

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PI

XX

DR WPI; 2001-488901/53.

DR

XX

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX

XX

PS Claim 27; SEQ ID No 20829; 487bp; English.

PS

XX

XX The present invention relates to human single exon nucleic acid probes

CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore

CC useful in grading and/or staging of diseases of the cervix, notably

CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SO Sequence 71 AA;

Query Match 67.9%; Score 38; DB 22; Length 71;

Best Local Similarity 77.8%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CISVPLSVP 9

Db 51 cisvplptp 59

RESULT 15

AAG30888

ID AAG30888 standard; Protein; 502 AA.

XX

AC AAG30888;

XX

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 37006.

XX

DE Protein identification: signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KM termination sequence.

XX

OS Arabidopsis thaliana.

OS

PN EP1033405-A2.

PN

XX

PD 06-SEP-2000.

PD

XX

PF 25-FEB-2000; 2000EP-0301439.

PF

XX

PR 25-FEB-1999; 99US-0121825.

PR

XX

PR 05-MAR-1999; 99US-0123180.

PR

XX

PR 09-MAR-1999; 99US-0123548.

PR

XX

PR 23-MAR-1999; 99US-0125788.

PR

XX

PR 25-MAR-1999; 99US-0126264.

PR

XX

PR 29-MAR-1999; 99US-0126785.

PR

XX

PR 01-APR-1999; 99US-0127462.

PR

XX

PR 06-APR-1999; 99US-0128234.

PR

XX

PR 08-APR-1999; 99US-0128714.

PR

XX

PR 16-APR-1999; 99US-0129845.

PR

XX

PR 19-APR-1999; 99US-0130077.

PR

XX

PR 21-APR-1999; 99US-0130449.

PR

XX

PR 23-APR-1999; 99US-0130510.

PR

XX

PR 28-APR-1999; 99US-0130891.

PR

XX

PR 30-APR-1999; 99US-0131449.

PR

XX

PR 04-MAY-1999; 99US-0132048.

PR

XX

PR 05-MAY-1999; 99US-0132407.

PR

XX

PR 06-MAY-1999; 99US-0132484.

PR

XX

PR 06-MAY-1999; 99US-0132485.

PR

XX

PR 06-MAY-1999; 99US-0132486.

PR

XX

PR 06-MAY-1999; 99US-0132487.

PR

XX

PR 07-MAY-1999; 99US-0132863.

PR

XX

PR 11-MAY-1999; 99US-0134256.

PR

XX

PR 14-MAY-1999; 99US-0134218.

PR

XX

PR 14-MAY-1999; 99US-0134219.

PR

XX

PR 14-MAY-1999; 99US-0134221.

PR

XX

PR 14-MAY-1999; 99US-0134370.

PR

XX

PR 18-MAY-1999; 99US-0134768.

PR

XX

PR 19-MAY-1999; 99US-0134941.

PR

XX

PR 20-MAY-1999; 99US-0135124.

PR

XX

PR 21-MAY-1999; 99US-0135353.

PR

XX

PR 24-MAY-1999; 99US-0135629.

PR

XX

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140359.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142399.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 9905-0162142.

Query Match 67.9%; Score 38; DB 21; length 502;

Best Local Similarity 85.7%; Pred. No. 2.3e+02;

Mismatches 0; Indels 0; Gaps 0;

QY 4 VPLSVPC 10

||||:||

DB 49 VPLSIPC 55

Search completed: June 17, 2002, 16:02:17
Job time: 426 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:56:39 ; Search time 75.98 Seconds
(without alignments)
3.215 Million cell updates/sec

Title: US-09-761-636A-13
Perfect score: 56
Sequence: 1 CISVPLSVPC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	64.3	9	1	US-08-331-383-33
2	36	64.3	9	1	US-08-549-008-43
3	36	64.3	9	1	US-08-802-981-143
4	34.5	61.6	72	3	US-08-722-719-58
5	34.5	61.6	73	3	US-08-722-719-26
6	34.5	61.6	73	3	US-08-722-719-57
7	34.5	61.6	76	3	US-08-722-719-55
8	34.5	61.6	76	3	US-08-722-719-24
9	34.5	61.6	76	3	US-08-722-719-56
10	34.5	61.6	77	3	US-08-722-719-59
11	34.5	61.6	77	3	US-08-722-719-22
12	34.5	61.6	77	3	US-08-722-719-61
13	34.5	61.6	77	4	US-09-044-796A-12
14	34.5	61.6	78	3	US-08-722-719-62
15	34.5	61.6	82	3	US-08-722-719-60
16	34.5	61.6	83	3	US-08-722-719-18
17	34.5	61.6	84	3	US-08-722-719-20
18	34.5	61.6	100	3	US-08-722-719-30
19	34.5	61.6	100	3	US-08-722-719-32
20	34.5	61.6	100	3	US-08-722-719-54
21	34.5	61.6	120	4	US-08-722-719-4
22	34.5	61.6	120	4	US-08-681-192-6
23	34.5	61.6	121	1	US-08-208-339A-2
24	34.5	61.6	137	2	US-08-421-144A-2
25	34	60.7	36	4	US-09-286-691-21
26	34	60.7	36	4	US-09-687-147-21
27	34	60.7	58	4	US-08-905-223-499

28	34	60.7	409	4	US-09-065-872-2	Sequence 2, Appl1
29	34	60.7	410	4	US-09-065-872-1	Sequence 1, Appl1
30	34	60.7	419	1	US-08-295-411-1	Sequence 1, Appl1
31	34	60.7	419	2	US-08-955-471-1	Sequence 1, Appl1
32	34	60.7	419	5	PCT-US92-10242-1	Sequence 1, Appl1
33	34	60.7	460	2	US-08-756-506-2	Sequence 2, Appl1
34	34	60.7	460	2	US-08-756-506-4	Sequence 2, Appl1
35	34	60.7	460	6	US-08-756-506-4	Sequence 4, Appl1
36	34	60.7	460	6	5270178-13	Patent No. 5270178
37	34	60.7	460	6	5270178-15	Patent No. 5270178
38	34	60.7	460	6	5270178-16	Patent No. 5270178
39	34	60.7	461	6	5225537-2	Patent No. 5225537
40	34	60.7	461	6	5270178-2	Patent No. 5270178
41	34	60.7	461	6	5270178-17	Patent No. 5270178
42	34	60.7	461	6	5270178-18	Patent No. 5270178
43	34	60.7	461	6	5460953-3	Patent No. 5460953
44	34	60.7	554	1	US-08-440-377A-5	Sequence 5, Appl1
45	34	60.7	554	2	US-08-687-852-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-331-383-33
Sequence 33, Application US/08331383
Patent No. 5605809
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
TITLE OF INVENTION: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of
NUMBER OF INVENTION: Proteases in Biological Samples and Methods and Use
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,383
FILING DATE: 28-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 16865-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-331-383-33

Query Match 64.3%, Score 36; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 3 SVPLSVPC 10
:::|::|

Db 2 AIPMSIPC 9

RESULT 2

US-08-549-008-43

Sequence 43, Application US/08549008

Patent No. 5714342

GENERAL INFORMATION:

APPLICANT: Komoriya, Akira

APPLICANT: Packard, Beverly S.

TITLE OF INVENTION: Compositions for the Detection of

TITLE OF INVENTION: Protease in Biological Samples and Methods of Use Thereof

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/549,008

FILING DATE: 27-OCT-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/331,383

FILING DATE: 28-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Ellen Lauver

REGISTRATION NUMBER: 32,762

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-549-008-43

Query Match

Best Local Similarity 50.0%; Score 36; DB 1; Length 9;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLSVPC 10

Db 2 AIPMSIPC 9

RESULT 3

US-08-802-981-143

Sequence 143, Application US/08802981

Patent No. 6037137

GENERAL INFORMATION:

APPLICANT: Komoriya, Akira

APPLICANT: Packard, Beverly S.

TITLE OF INVENTION: Compositions for the Detection of Enzyme

TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/802,981

FILING DATE: 20-FEB-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 016865-000300US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 143:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-802-981-143

Query Match

Best Local Similarity 50.0%; Score 36; DB 3; Length 9;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLSVPC 10

Db 2 AIPMSIPC 9

RESULT 4

US-08-722-719-58

Sequence 58, Application US/08722719

Patent No. 6001606

GENERAL INFORMATION:

APPLICANT: ROSEN, CRAIG A.

APPLICANT: RUBIN, STEVEN M.

APPLICANT: LI, HAODONG

APPLICANT: ADAMS, MARK D.

TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITOR

TITLE OF INVENTION: FACTOR-1 (MP1F-1), MONOCYTE COLONY INHIBITORY FACTOR

TITLE OF INVENTION: (M-CIF), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/722,719

FILING DATE: 30-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,209

FILING DATE: 22-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/208,339

FILING DATE: 08-MAR-1994

APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-722-719-57

Query Match 61.6%; Score 34.5; DB 3; Length 73;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 CIS-VPLSVPC 10
||| | :||
Db 8 CISTYPRSI PC 18

RESULT 7
US-08-722-719-55
Sequence 55, Application US/08722719
Patent No. 6001606
GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MIF-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-CIF), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-722-719-55

Query Match 61.6%; Score 34.5; DB 3; Length 75;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 CIS-VPLSVPC 10
||| | :||
Db 10 CISTYPRSI PC 20

RESULT 8
US-08-722-719-24
Sequence 24, Application US/08722719
Patent No. 6001606
GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MIF-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-CIF), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775

FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-722-719-24

Query Match 61.6%; Score 34.5; DB 3; Length 76;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 CIS-VPLSVC 10
111 1 1:11
DB 11 CISTPRSPIC 21

RESULT 9
US-08-722-719-56
Sequence 56, Application US/08722719
Patent No. 6001606

GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG M.
APPLICANT: ADAMS, MARK D.
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MPLF-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-CIF), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688

NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-722-719-56

Query Match 61.6%; Score 34.5; DB 3; Length 76;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 CIS-VPLSVC 10
111 1 1:11
DB 11 CISTPRSPIC 21

RESULT 10
US-08-722-719-59
Sequence 59, Application US/08722719
Patent No. 6001606

GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG M.
APPLICANT: ADAMS, MARK D.
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MPLF-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-CIF), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-722-719-59

Query Match 61.6%; Score 34.5; DB 3; Length 76;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 CIS-VPLSVC 10
111 1:11
DB 11 CISTYPRSPIC 21

RESULT 11
US-08-722-719-22
Sequence 22, Application US/08722719
Patent No. 6001606
GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG
APPLICANT: ADAMS, MARK D.
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MP1F-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-C1F), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-722-719-22

Query Match 61.6%; Score 34.5; DB 3; Length 77;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 CIS-VPLSVC 10
111 1:11
DB 12 CISTYPRSPIC 22

RESULT 12
US-08-722-719-61
Sequence 61, Application US/08722719
Patent No. 6001606
GENERAL INFORMATION:
APPLICANT: ROSEN, CRAIG A.
APPLICANT: RUBIN, STEVEN M.
APPLICANT: LI, HAODONG
APPLICANT: ADAMS, MARK D.
TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS AND METHODS FOR
TREATING DISEASE STATES WITH MYELOID PROGENITOR INHIBITORY
TITLE OF INVENTION: FACTOR-1 (MP1F-1), MONOCYTE COLONY INHIBITORY FACTOR
TITLE OF INVENTION: (M-C1F), AND MACROPHAGE INHIBITORY FACTOR-4 (MIP-4)
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,719
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,209
FILING DATE: 22-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/208,339
FILING DATE: 08-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,881
FILING DATE: 05-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,682
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,775
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEERE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0330007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-722-719-60

Query Match 61.6%; Score 34.5; DB 3; Length 82;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 CIS-VPLSVPC 10
Db 17 CISTYPRISIPC 27

Search completed: June 17, 2002, 15:56:40
Job time: 89 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 15:58:27 ; Search time 95.43 seconds
(without alignments)
9.062 Million cell updates/sec

Title: US-09-761-636A-14

Perfect score: 52

Sequence: 1 CISVPLVPC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	75.0	297	2 T25817	hypothetical prote
2	39	75.0	347	2 A81988	probable N-acetyl-
3	39	75.0	347	2 D81043	N-acetyl-gamma-glu
4	39	75.0	416	2 A61404	keratin A, type I,
5	39	75.0	416	2 A46559	keratin, type I, h
6	38	73.1	154	2 D82795	phage-related endo
7	38	73.1	577	2 T00416	hypothetical prote
8	38	73.1	1015	2 C84918	probable ATP-depen
9	38	73.1	3623	2 T08618	intrinsic factor-B
10	37	71.2	287	2 H64690	type IIS restricti
11	37	71.2	407	2 T11985	hypothetical prote
12	36	69.2	1639	2 T50119	probable sensory t
13	35	67.3	61	2 J01086	nodulin-14 precurs
14	35	67.3	278	2 AF0936	probable Arac-fam
15	35	67.3	305	2 A82728	conserved hypotet
16	35	67.3	496	2 S51668	tyrosine kinase
17	35	67.3	1210	2 D88013	protein K10B4.1 (I
18	35	67.3	1440	1 SYH0UT	multifunctional am
19	35	67.3	2240	1 T37057	probable multi-dom
20	34	65.4	397	2 AC1877	hypothetical prote
21	34	65.4	551	2 JC7562	globlastoma RING
22	34	65.4	1876	2 T13801	phosphoinositide 3
23	33	63.5	339	2 A13609	daunorubicin resis
24	33	63.5	370	2 T25627	hypothetical prote
25	33	63.5	434	2 T01390	hypothetical prote
26	33	63.5	491	2 T46915	hypothetical prote
27	33	63.5	497	1 S53834	NADH dehydrogenase
28	33	63.5	498	2 T45465	acetate CoA-transf
29	33	63.5	725	2 JE0099	neural cell adhesi

30	33	63.5	771	2 F70757	probable cation tr
31	33	63.5	956	2 A65072	hypothetical prote
32	33	63.5	956	2 B91098	probable dehydroge
33	33	63.5	956	2 F85943	probable dehydroge
34	33	63.5	957	2 T41246	probable ribonucle
35	33	63.5	1021	2 S26985	probable DNA-direc
36	33	63.5	1046	2 A26838	prestalk protein p
37	33	63.5	1253	1 VHWV	structural polypro
38	33	63.5	1827	2 A35694	cuti protein - fis
39	33	63.5	1828	2 T41455	cuti protein - fis
40	32	61.5	58	2 A23101	lon protein - Esch
41	32	61.5	72	1 W5ML39	E5 protein - human
42	32	61.5	84	2 A61594	thrombospondin - b
43	32	61.5	115	2 D71541	probable inclusion
44	32	61.5	174	2 T02683	hypothetical prote
45	32	61.5	185	1 A47196	dual specificity p

ALIGNMENTS

RESULT 1

T25817
hypothetical protein K12D9.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T25817

R:Graves, T.
submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans cosmid K12D9.

A:Reference number: 22093

A/Accession: T25817

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-297 <GRAB>

A/Cross-references: EMBL:U80030; PIDN:AAB37600.1; GSPDB:GN00023; CESP:K12D9.3

A/Experimental source: Strain Bristol N2; Clone K12D9

C/Genetics:

A:Gene: CESP:K12D9.3

A/Map position: 5

A/Introns: 165/3; 208/3

Query Match 75.0%; Score 39; DB 2; Length 297;

Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CISVPLVPC 9

Db 169 CISPLVPC 177

RESULT 2

A81988
probable N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) NMA0676 [Import

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C/Accession: A81988

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.;

Nature 404, 502-506, 2000

A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22

A/Reference number: A81775; MUID:20222556

A/Accession: A81988

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-347 <PAR>

A/Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83963.1; PID:g

A/Experimental source: serogroup A, strain Z2491

C/Genetics:

A:Gene: argC; NMA0676

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

C/Keywords: oxidoreductase

Query Match
Best Local Similarity 75.0%; Score 39; DB 2; Length 347;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVPLVP 8
1:1:1111
Db 156 C1SVPLVP 163

RESULT 3

N-acetyl-gamma-glutamyl-phosphate reductase NMB1787 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: DB1043

R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755

A:Accession: DB1043
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-347 <TRN>

A:Cross-references: GB:AE002528; GB:AE002098; NID:g7227034; PIDN:AAF42126.1; PID:g722704
A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1787

C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase

Query Match
Best Local Similarity 75.0%; Score 39; DB 2; Length 347;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVPLVP 8
1:1:1111
Db 156 C1SVPLVP 163

RESULT 4

Keratin A, type I - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 03-May-1996
C:Accession: A61404

R:Kayes, P.S.; McNab, A.R.; Rea, T.J.; Groppi, V.; Kawabe, T.T.; Buhl, A.E.; Bertolino,
J. Invest. Dermatol. 97, 835-842, 1991
A:Title: Hair-specific keratins: characterization and expression of a mouse type I kerat
A:Reference number: A61404; MUID:92013235

A:Accession: A61404

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-416 <KAY>

C:Superfamily: cytoskeletal keratin

Query Match
Best Local Similarity 75.0%; Score 39; DB 2; Length 416;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 C1SVPLVP 9
1:1:1111
Db 385 C1SVPLVP 393

RESULT 5

keratin, type I, hair (clone MHKA-1) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 13-Aug-1999

C:Accession: A46559

R:Bertolino, A.P.; Checkla, D.M.; Notterman, R.; Sklaver, I.; Schliff, T.A.; Freedbu

J. Invest. Dermatol. 91, 541-546, 1988

A:Title: Cloning and characterization of a mouse type I hair keratin cDNA.

A:Reference number: A46559; MUID:89054860

A:Accession: A46559

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <BBR>

A:Cross-references: GB:M27734; NID:9198585; PIDN:AAA93972.1; PID:9198586

C:Superfamily: cytoskeletal keratin

Query Match
Best Local Similarity 75.0%; Score 39; DB 2; Length 416;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 C1SVPLVP 9
1:1:1111
Db 385 C1SVPLVP 393

RESULT 6

phage-related endolysin XF0513 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 23-Mar-2001

C:Accession: DB2795

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide s

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: DB2795

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-154 <SIM>

A:Cross-references: GB:AE003900; GB:AE003849; NID:9105366; PIDN:AAF83323.1; GSPDB:

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinech, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carr

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.;

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.;

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; da Si

M.; Tshahro, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0513

C:Superfamily: phage T4 lysozyme; phage T4 lysozyme homology

Query Match
Best Local Similarity 73.1%; Score 38; DB 2; Length 154;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVPLVP 8
1:1:1111
Db 78 C1SVPLVP 85

RESULT 7

hypothetical protein T30B22.1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 24-Nov-1999
C:Accession: T00416
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Masc
submitted to the EMBL Data Library, October 1998
A:Description: Arabidopsis thaliana chromosome II BAC T30B22 genomic sequence.
A:Reference number: 214149
A:Accession: T00416
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-577 <R0U>
A:Cross-references: EMBL:AC002535; NID:g2529657; PID:g2529658
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 2
A:introns: 87/3; 126/3; 175/1; 264/3; 329/3; 507/3
C:Superfamily: Arabidopsis thaliana hypothetical protein T30B22.1

Query Match 73.1%; Score 38; DB 2; Length 577;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVPLVP 8
|:|||||
Db 238 C1SVPLVP 245

RESULT 8
C84918
probable ATP-dependent RNA helicase A [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84918
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: C84918
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1015 <STO>
A:Cross-references: GB:AE002093; NID:g3738282; PIDN:AAC63624.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g47680
A:Map position: 2

Query Match 73.1%; Score 38; DB 2; Length 1015;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 C1SVPLVP 8
|:|||||
Db 676 C1SVPLVP 683

RESULT 9
T08618
intrinsic factor-B12 receptor CUBILIN precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08618
R:Koeztup, S.K.; Kozyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, H.H.; Brault,
J. Biol. Chem. 273, 5235-5242, 1998
A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies
A:Reference number: 216459; MUID:98148073
A:Accession: T08618
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <MOE>
A:Cross-references: EMBL:AF022247; NID:g3834379; PIDN:AACT1661.1; PID:g3834380

C:Genetics:
A:Gene: CUBILIN
C:Superfamily: unassigned EGF-related proteins: EGF homology
C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral mem
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
F:133-164/Domain: EGF homology <EGF1>
F:436-467/Domain: EGF homology <EGF>

Query Match 73.1%; Score 38; DB 2; Length 3623;
Best Local Similarity 66.7%; Pred. No. 2,3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9
|:|||||
Db 313 C1SVPLVPC 321

RESULT 10
H64690
type IIS restriction enzyme M2 protein - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 18-Jun-1999
C:Accession: H64690
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischman
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; P
son, J.P.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wat
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fra
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: H64690
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-287 <TON>
A:Cross-references: GB:AE000637; GB:AE000511; NID:g2314536; PIDN:AD08412.1; PID:
C:Genetics:
A:Start codon: TTG
C:Superfamily: site-specific DNA-methyltransferase (adenine-specific) HpaI

Query Match 71.2%; Score 37; DB 2; Length 287;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 C1SVPLVP 8
|:|||||
Db 86 C1SVPLVP 93

RESULT 11
T19895
hypothetical protein C41G7.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19895
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19192
A:Accession: T19895
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-407 <WIL>
A:Cross-references: EMBL:281048; PIDN:CA02841.1; GSPDB:GN00019; CESP:C41G7.4
A:Experimental source: clone C41G7
C:Genetics:
A:Gene: CESP:C41G7.4
A:Map position: 1
A:introns: 27/2; 107/3; 151/3; 333/3; 373/3

Query Match 71.2%; Score 37; DB 2; Length 407;

Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 C1SVPLVPC 9
| | | | |
Db 113 CSSNPLVPC 121

RESULT 12

750119
Probable sensory transduction histidine kinase [Imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50119
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, February 2000
A:Reference number: 225039
A:Accession: T50119
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1639 <SEE>
A:Cross-references: EMBL:AL157734; PIDN:CAB75776.1; GSPDB:GN00066; SPDB:SPAC1834.08
A:Experimental source: strain 972h(-); cosmid c1834
C:Genetics:
A:Gene: SPDB:SPAC1834.08
A:Map position: 1
A:Introns: 23/3

Query Match

Best Local Similarity 69.2%; Score 36; DB 2; Length 1639;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLVPC 9
| | | | |
Db 557 SVPLVPC 563

RESULT 13

QJ01086
nodulin-14 precursor - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C>Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 20-Apr-2001
C:Accession: QJ01086
R:Scheres, B.; van Engelen, F.; van der Knaap, E.; van de Wiel, C.; van Kammen, A.; Biss Plant Cell 2, 687-700, 1990
A:Title: Sequential induction of nodulin gene expression in the developing pea nodule.
A:Reference number: QJ01084; MUID:93005665
A:Accession: QJ01086
A:Molecule type: mRNA
A:Residues: 1-61 <SCH>
A:Cross-references: GB:S45162; NID:9257036; PIDN:AAE23538.1; PID:9257037
A:Experimental source: root nodule
C:Comment: The gene encoding this protein is expressed in the second stage of root nodulation.
C:Genetics:
A:Gene: ENOD14
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-61/Product: nodulin 14 #status predicted <MAT>

Query Match

Best Local Similarity 71.4%; Score 35; DB 2; Length 61;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SVPLVPC 9
| | | | |
Db 21 NIPLVPC 27

RESULT 14

AF0936
Probable AraC-family transcription regulator STY3756 [Imported] - Salmonella enterica subsp. *Salmonella enterica* subsp. *enterica* serovar Typhi
C:Species: Salmonella enterica subsp. enterica serovar Typhi

A>Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AF0936
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Connerly, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; F.

'S.; Moulé, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Steven
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica*
A:Reference number: AB0502; PMID:11677608
A:Accession: AF0936
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-278 <PAR>
A:Cross-references: GB:AL153382; PIDN:CAD09512.1; PID:q16504629; GSPDB:GN00176
C:Genetics:
A:Gene: STY3756

Query Match

Best Local Similarity 67.3%; Score 35; DB 2; Length 278;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISVPLVPC 9
| | | | |
Db 56 ITAPLVPC 63

RESULT 15

AB2728
conserved hypothetical protein XF1068 [Imported] - *Xylella fastidiosa* (strain 9a5c)
C:Species: *Xylella fastidiosa*
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: AB2728
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A:Reference number: AB2515; MUID:20365717
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: AB2728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <STM>
A:Cross-references: GB:AE003943; GB:AE003849; NID:9106006; PIDN:AAE83878.1; GSPDB:
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
Briñon, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
as-Neto, E.; Docena, C.; El-Dorry, H.; Feliciani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferio, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.;
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.;
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki,
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmiter
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Si
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1068

Query Match

Best Local Similarity 57.1%; Score 35; DB 2; Length 305;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

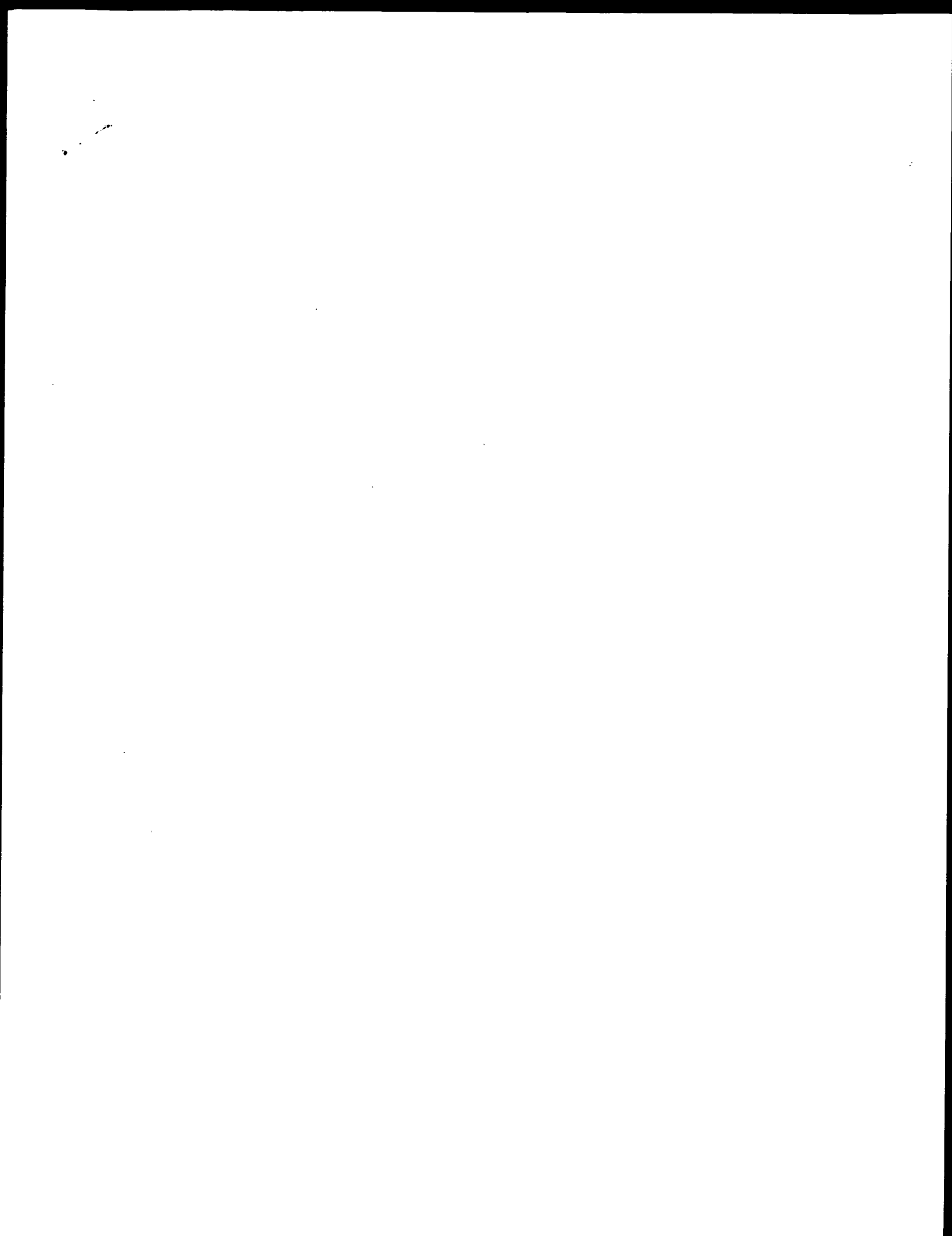
QY 3 SVPLVPC 9
| | | | |
Db 3 SIPLVPC 9

Mon Jun 17 16:11:25 2002

us-09-761-636a-14.open.rpr

Page 5

Search completed: June 17, 2002, 15:58:28
Job time: 197 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:24:32 ; Search time 44.84 Seconds

7.772 Million cell updates/sec

Title:	US-09-761-636A-14
Perfect score:	52
Sequence:	1 CISVPLVPC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	75.0	416	1	K1M1_MOUSE	Q61765 mus musculus
2	35	67.3	61	1	N014_PEA	P26415 pisum sativ
3	35	67.3	78	1	VE5_HPV70	P50774 human papill
4	35	67.3	279	1	MOV6_CMVAS	O36639 cucumber mo
5	35	67.3	490	1	FA10_RABIT	O19045 oryctolagus
6	35	67.3	1108	1	EMBA_MYCAV	P71485 mycobacteri
7	35	67.3	1440	1	SYEP_HUMAN	P07814 homo sapien
8	34	65.4	551	1	RN27_HUMAN	O99zr9 homo sapien
9	34	65.4	551	1	RN27_MOUSE	O99pj2 mus musculu
10	34	65.4	966	1	M172_HUMAN	Q14596 homo sapien
11	34	65.4	1173	1	TSP1_XENIA	P33448 xenopus lae
12	33	63.5	497	1	N14M_ACACA	Q37375 acanthamoeb
13	33	63.5	771	1	CTPG_MYCTU	O10866 mycobacteri
14	33	63.5	1021	1	DPM6_NEUCR	P33557 neurospora
15	33	63.5	1046	1	PS7A_DICDI	P11976 dictyostell
16	33	63.5	1253	1	POL5_SEV	P03315 semliki for
17	33	63.5	1828	1	CUT11_SCHPO	P18296 schizosach
18	32	61.5	72	1	VE5_HPV39	P24833 human papill
19	32	61.5	185	1	DUS3_HUMAN	P51452 homo sapien
20	32	61.5	185	1	DUS3_MOUSE	O947x3 mus musculu
21	32	61.5	262	1	PLC1_CAEEL	O95841 caenorhabd
22	32	61.5	305	1	CORA_HPBDO	P03154 duck hepati
23	32	61.5	312	1	NRT1_CHICK	P55806 gallus gallu
24	32	61.5	346	1	GATD_ECOLI	P37190 escherichia
25	32	61.5	387	1	DHB2_HUMAN	P37059 homo sapien
26	32	61.5	413	1	GAT1_MOUSE	P17679 mus musculu
27	32	61.5	413	1	GAT1_RAT	P43429 rattus norv
28	32	61.5	459	1	RSP6_CHLRE	Q01657 chlamydomo
29	32	61.5	467	1	K1M6_HUMAN	P56013 homo sapien
30	32	61.5	662	1	T9S2_MOUSE	O78023 mus musculu
31	32	61.5	663	1	T9S2_HUMAN	O99805 homo sapien
32	32	61.5	824	1	DP20_HUMAN	O99u16 homo sapien
33	32	61.5	1170	1	TSP1_BOVIN	Q28178 bos taurus

34	32	61.5	1170	1	TSP1_HUMAN	P07996	homo sapien
35	32	61.5	1170	1	TSP1_MOUSE	P33441	mus musculus
36	31.5	60.6	569	1	MANC_PIRSP	P55298	ptarmigan s
37	31.5	60.6	571	1	MANB_PIRSP	P55297	ptarmigan s
38	31.5	60.6	669	1	COTE_HUMAN	P81408	homo sapien
39	31.5	60.6	874	1	STYM_SCHPO	P09828	schizosacch
40	31	59.6	144	1	CNI_DROME	P40888	drosophila
41	31	59.6	210	1	CAC2_HAECO	P12525	haemophilus
42	31	59.6	284	1	CX32_MICNO	P19125	microgonius
43	31	59.6	288	1	TYVE_METTA	O56414	methanococc
44	31	59.6	347	1	ZEP2_MOUSE	P08043	mus musculus
45	31	59.6	497	1	CEP2_RABIT	P22687	oryctolagus

ALIGNMENTS

RESULT	1			
ID	KIM1_MOUSE	STANDARD;	PR1;	416 AA.
AC	061765;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Keratin, type I cuticular HAI (hair keratin, type I HAI) (HKA-1).			
GN	KRTHAI OR KRT1-1 OR KHA1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RA	[1]			
RL	SEQUENCE FROM N.A.			
RC	STRATN-C57BL/6J; TISSUE=Hair;			
RX	MIDDLELINE=89054860; PubMed=2461417;			
RA	Bertolini A.P., Checkla D.M., Notterman R., Sklaver I., Schliff T.A.,			
RA	Frederberg I.M., Didona G.J.;			
RT	Cloning and characterization of a mouse type I hair keratin cDNA.;			
CC	J. Invest. Dermatol. 91:541-546(1988).			
CC	-I- MISCELLANEOUS: THERE ARE TWO TYPES OF HAIR/MICROFIBRILLAR KERATIN,			
CC	I (ACIDIC) AND II (NEUTRAL TO BASIC).			
CC	-I- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M27734; AAA39372.1; -.			
DR	MGD; MG1:1309993; Krt1-1.			
DR	InterPro; IPR001664; IF.			
DR	InterPro; IPR002957; Keratin_I.			
DR	Pfam; PF000038; filament_1.			
DR	PRINTS; PR01248; TYPE1KERATIN.			
DR	PROSITE; PS00226; IF; 1.			
KW	Intermediate filament; Coiled coil; Keratin.			
FT	DOMAIN 2			
FT	DOMAIN 57			
FT	DOMAIN 364			
FT	DOMAIN 57			
FT	DOMAIN 92			
FT	DOMAIN 103			
FT	DOMAIN 204			
FT	DOMAIN 220			
FT	DOMAIN 305			
FT	SITE 305			
SQ	SEQUENCE 416 AA; 47112 MW; 9C3759CF4E3A3BF3 CRC64;			

Query Match	75.0%;	Score 39;	DB 1;	Length 416;
Best Similarity	66.7%;	Pred. No. 5.8;		
Best Local Matches	6;	Conservative 1;	Mismatches 2;	Indels 0;
				Gaps 0;

OY 1 C1SVPLVPC 9
Db 385 C1SVNCPVPC 393

RESULT 2

NO14_PEA
ID NO14_PEA STANDARD: PRT: 61 AA.

DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Nodulin 14 precursor (N-14) (Fragment).
GN ENOD14.

OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=CV, SPARKLE: TISSUE=Root nodules;
RX MEDLINE=93005655; PubMed=2152123;

RA Scheres B., van Engelen F., van der Knaap E., van de Wiel C.,
van Kammen A., Bisseling T.;
RT "Sequential induction of nodulin gene expression in the developing
RL pea nodule.";
RT Plant Cell 2:687-700(1990).

CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE SECOND STAGE OF ROOT NODULE
FORMATION.
CC -1- MISCELLANEOUS: CONTAINS 4 CYSTEINES ARRANGED IN TWO PAIRS IN SUCH
A WAY THAT THEY MIGHT BE CAPABLE OF BINDING A METAL ION.
CC -1- SIMILARITY: TO NODULIN 3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: S45162; AAB23538.1; -
DR PIR: JQ1085; JQ1085.
DR PIR: JQ1086; JQ1086.
KW Modulation; signal; Metal-binding.
FT NON_TER 1 1
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 61 NODULIN 14.
SQ SEQUENCE 61 AA: 7175 MW; D508059695F628B8 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 61;
Best Local Similarity 71.4%; Pred. No. 4.9;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 SVPLVPC 9
Db 21 N1PLVPC 27

RESULT 3

VE3_HP70
ID VE3_HP70 STANDARD: PRT: 78 AA.

AC P50774;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable E5 protein.
GN E5.

OS Human papillomavirus type 70.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.
OX NCBI_TaxID=39457;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=96249586; PubMed=8815087;
RA Forslund O., Hansson B.G.;

RT "Human papillomavirus type 70 genome cloned from overlapping PCR
RT products: complete nucleotide sequence and genomic organization.";
RL J. Clin. Microbiol. 34:802-809(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U21941; AAC54855.1; -
DR InterPro: IPR004270; Papilloma_E5.
DR Pfam: PF03025; Papilloma_E5; 1.
KW Early protein.
SQ SEQUENCE 78 AA: 9262 MW; DFEA59CB35042D1 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 78;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 8
Db 24 C1SVPLVPC 31

RESULT 4
ID MOV_P_CNVAS STANDARD: PRT: 279 AA.

AC O36639;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cell-to-cell movement protein (MP) (3A protein).

GN 3A.
OS Cucurbit mosaic virus (strain As) (CMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;

OC Cucumovirus.
OX NCBI_TaxID=117118;

RP SEQUENCE FROM N.A.
RA Kim S.H., Park Y.I.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN TRANSPORT OF THE VIRUS FROM THE INITIALLY
CC INFECTED CELLS TO ADJACENT CELLS.

CC -1- SIMILARITY: BELONGS TO THE CUCUMOVIRUS/BROMOVIRUS 3A FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AF013291; AAB67166.1; -
DR InterPro: IPR000603; 3A_mov.
DR Pfam: PF00803; 3A; 1.
KW DNA-binding; Transport.
SQ SEQUENCE 279 AA: 30509 MW; 2B4705928DC1069 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 279;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLVPC 9
1:111 11
DB 85 CYSVPLVPC 93

RESULT 5
FA10. RABIT STANDARD; PRT: 490 AA.
ID FA10. RABIT
AC 019045:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN F10.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97256311; PubMed=9101642;
RA Pendurthi U.R., Anderson K.D., James H.L.;
RT Characterization of a full-length cDNA for rabbit factor X.";
RL Thromb. Res. 85:503-514(1997).
CC -1- FUNCTION: FACTOR XA IS A VITAMIN K-DEPENDENT GLYCOPROTEIN THAT
CC CONVERTS PROTHROMBIN TO THROMBIN IN THE PRESENCE OF FACTOR VA,
CC C⁺⁺, AND PHOSPHOLIPID DURING BLOOD CLOTTING.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-1-thr and then
CC Arg-1-ile bonds in prothrombin to form thrombin.
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
CC MORE DISULFIDE BONDS.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC CALCIUM (BY SIMILARITY).
CC -1- PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC (BY SIMILARITY).
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC ANOTHER SITE, BEYOND THE GLA DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF003200; AAB62342.1; -
DR HSSP: P00742; IHCG.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002383; GLA_blood.
DR InterPro: IPR001254; Trypsin.
DR InterPro: IPR000294; Vitk_dep_GLA.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00594; gla_1.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00001; GLABLOOD.
DR SMART: SM00179; EGF_CA_1.
DR SMART: SM00001; EGF_like_1.
DR SMART: SM00069; gla_1.
DR SMART: SM00020; tryp_spc_1.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA_1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Glycoprotein: Hydrolase: Serine protease; Plasma: Blood coagulation;
KW gamma-carboxyglutamic acid: Hydroxylation; Calcium-binding; Vitamin K;
KW Signal; Zymogen; EGF-like domain; Repeat.
FT SIGNAL 1 20
FT PROPEP 21 40
FT CHAIN 41 180
FT CHAIN 184 490
FT PROPEP 184 232
FT CHAIN 233 490
FT DOMAIN 86 122
FT DOMAIN 125 165
FT DOMAIN 233 490
FT MOD_RES 46 46
FT MOD_RES 47 47
FT MOD_RES 54 54
FT MOD_RES 54 54
FT MOD_RES 56 56
FT MOD_RES 59 59
FT MOD_RES 60 60
FT MOD_RES 65 65
FT MOD_RES 66 66
FT MOD_RES 69 69
FT MOD_RES 72 72
FT MOD_RES 75 75
FT MOD_RES 79 79
FT MOD_RES 103 103
FT ACT_SITE 274 274
FT ACT_SITE 320 320
FT ACT_SITE 417 417
FT DISULFID 90 101
FT DISULFID 112 121
FT DISULFID 129 140
FT DISULFID 136 149
FT DISULFID 151 164
FT DISULFID 172 340
FT DISULFID 239 244
FT DISULFID 259 275
FT DISULFID 388 402
FT DISULFID 413 441
FT CARBOHYD 61 61
FT CARBOHYD 187 187
FT CARBOHYD 205 205
SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85AFA2A6D11 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 490;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CISVPLVPC 9
DB 164 CISTELFPC 172

```

RESULT      6
EMBL_MYCAV  STANDARD; PRT: 1108 AA.
ID   P71485;
AC   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DE   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Probable arabinosyltransferase A (EC 2.4.2.-).
GN   EMBA.
OS   Mycobacterium avium.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1764;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2151;
RX   MEDLINE=97030297; PubMed=8876238;
RA   Belanger A.E., Besta G.S., Ford M.E., Mikusova K., Belisle J.T.,
RA   Brennan P.J., Inamine J.M.;
RT   "The embA genes of Mycobacterium avium encode an arabinosyl
RT   transferase involved in cell wall arabinan biosynthesis that is the
RT   target for the antitubercular drug ethambutol";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:11919-11924(1996).
CC   -1- FUNCTION: Arabinosyl transferase responsible for the
CC   polymerization of arabinose into the arabinan of arabinogalactan.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -1- SIMILARITY: BELONGS TO THE EMB FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: U06560; AAC4547.1; -
KM   Transferase; Glycosyltransferase; Transmembrane; Cell wall;
KW   Antibiotic resistance.
FT   TRANSMEM  12
FT   TRANSMEM  34
FT   TRANSMEM  204 223 POTENTIAL.
FT   TRANSMEM  258 280 POTENTIAL.
FT   TRANSMEM  334 356 POTENTIAL.
FT   TRANSMEM  368 387 POTENTIAL.
FT   TRANSMEM  397 414 POTENTIAL.
FT   TRANSMEM  421 443 POTENTIAL.
FT   TRANSMEM  463 482 POTENTIAL.
FT   TRANSMEM  531 553 POTENTIAL.
FT   TRANSMEM  582 604 POTENTIAL.
FT   TRANSMEM  616 638 POTENTIAL.
FT   TRANSMEM  653 675 POTENTIAL.
FT   TRANSMEM  696 718 POTENTIAL.
SQ   SEQUENCE 1108 AA; 117404 MW; 0ED8E648768075FA CRC64;

```

```

Query Match      67.3%; Score 35; DB 1; Length 1108;
Best Local Similarity 62.5%; Pred. No. 80;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 C1SVPLVP 8
    1::1111:1
DB 28 CLAVPLVP 35

```

```

RESULT      7
SYEP_HUMAN  STANDARD; PRT: 1440 AA.
ID   P07814;
AC   01-AUG-1988 (Rel. 08, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bifunctional aminoacyl-tRNA synthetase [includes: Glutamyl-tRNA

```

```

DE synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase); Prolyl-tRNA
DE synthetase (EC 6.1.1.15) (Proline--tRNA ligase)].
GN EPRS OR QPRS OR GLNS OR PARS.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9107633; PubMed=1988429;
RA   Felt R., Knippers R.;
RT   "The primary structure of human glutamyl-tRNA synthetase. A highly
RT   conserved core, amino acid repeat regions, and homologies with
RT   translation elongation factors.";
RL   J. Biol. Chem. 266:1448-1455(1991).
RN   [2]
RP   PRELIMINARY SEQUENCE OF 96-887 FROM N.A.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=8826251; PubMed=3290852;
RA   Thoenes P., Felt R., Schray B., Kunze N., Knippers R.;
RT   "The core region of human glutamyl-tRNA synthetase homologues with
RT   the Escherichia coli and yeast enzymes.";
RL   Nucleic Acids Res. 16:5391-5406(1988).
CC   -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC   diphosphate + L-glutamyl-tRNA(Glu).
CC   -1- CATALYTIC ACTIVITY: ATP + L-proline + tRNA(Pro) = AMP +
CC   diphosphate + L-prolyl-tRNA(Pro).
CC   -1- SUBUNIT: COMPONENT OF THE MULTISUBUNIT COMPLEX WHICH IS
CC   COMPRISED OF A BIFUNCTIONAL GLUTAMYL-PROLYL-TRNA SYNTHETASE, THE
CC   MONOSPECIFIC ISOLEUCYL, LEUCYL, GLUTAMINYLYL, METHIONYL, LYSYL,
CC   ARGINYLYL, AND ASPARYLYL-TRNA SYNTHETASES AS WELL AS THREE AUXILIARY
CC   PROTEINS, P18, P48 AND P43.
CC   -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CLASS-I
CC   AMINOACYL-TRNA SYNTHETASE FAMILY.
CC   -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CLASS-II
CC   AMINOACYL-TRNA SYNTHETASE FAMILY.
CC   -1- SIMILARITY: CONTAINS 3 "WHEP-TRS" DOMAINS.
CC   -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE A GLUTAMINYLYL-TRNA
CC   SYNTHETASE.
CC   -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A NUMBER OF
CC   SEQUENCE ERRORS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: X54326; CAA38224.1; -
DR   EMBL: X07466; CAA30354.1; ALT_SEQ.
DR   PIR: A38663; SYHUOT.
DR   HSSP: P00962; IGTR.
DR   MIM: 138295; -
DR   InterPro: IPR002106; AA_tRNA_ligase-II.
DR   InterPro: IPR004046; GST_C.
DR   InterPro: IPR004154; HGTP_antlicodon.
DR   InterPro: IPR000738; WHEP-TRS.
DR   InterPro: IPR000924; tRNA-synt_1c.
DR   InterPro: IPR002314; tRNA-synt_1c.
DR   InterPro: IPR001412; tRNA-synt_1.
DR   InterPro: IPR002316; tRNA-synt_pro.
DR   Pfam: PF00043; GST_C. 1.
DR   Pfam: PF03129; HGTP_antlicodon. 1.
DR   Pfam: PF00749; tRNA-synt_1c. 1.
DR   Pfam: PF00587; tRNA-synt_2b. 1.
DR   Pfam: PF00458; WHEP-TRS. 3.
DR   PRINTS: PR00987; TRNASYNTHGLD.
DR   PRINTS: PR01046; TRNASYNTHPRO.
DR   PROSITE: PS00178; AA_tRNA_LIGASE_I. 1.
DR   PROSITE: PS00179; AA_tRNA_LIGASE_II_1; FALSE_NEG.
DR   PROSITE: PS00339; AA_tRNA_LIGASE_II_2; FALSE_NEG.

```



```

CC or send an email to license@lsb-stb.ch).
CC -----
DR EMBL: AF281047; AAG53088.1;
DR EMBL: AF220035; AAG53489.1;
DR MGD: MGI:193302; Trlrm8
DR InterPro: IPR000315; Znf_dbox.
DR InterPro: IPR001841; Znf_finger.
DR Pfam: PF00643; zf-B_box; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00119; ZF_BOX; FALSE_NG.
DR PROSITE: PS00518; ZF_RING; 1.
DR PROSITE: PS00089; ZF_RING_2; 1.
KW Nuclear protein; Coiled coil; Zinc-finger; Repeat.
FT ZN_FING 15 56 RING-TYPE.
FT ZN_FING 92 132 B_BOX-TYPE 1.
FT ZN_FING 140 182 B_BOX-TYPE 2.
FT DOMAIN 181 249 COILED COIL (POTENTIAL).
FT DOMAIN 274 295 COILED COIL (POTENTIAL).
FT CONFLICT 259 259 G -> A (IN REF. 2).
SQ SEQUENCE 551 AA; 61591 MW; CECB2D70C8B1E0B6 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 551;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLVPC 9
Db 361 SVPLVPC 367

RESULT 10
M172_HUMAN STANDARD; PRT; 966 AA.
ID M172_HUMAN
AC Q14596; Q13173; Q15026;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Membrane component, chromosome 17, surface marker 2 (Ovarian carcinoma
DE antigen CA125) (1A1-3B) (next to BRCA1 gene 1 protein).
OS Homo sapiens (Human).
GN M17S2 OR 1A13B OR NBRI OR KIAA0049.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RX MEDLINE=94348506; PubMed=8069304;
RA Campbell I.G., Nicolai H.M., Foulkes W.D., Senger G., Stamp G.W.,
RA Allan G., Boyer C., Jones K., Bast R.C. Jr., Solomon E., Trowsdale J.,
RA Black D.M.,
RT "A novel gene encoding a B-box protein within the BRCA1 region at
RT 17q21.1".
RL Hum. Mol. Genet. 3:589-594(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarayashi Y., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (K1AA0041-K1AA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
RN [3]
RP SEQUENCE OF 412-508 FROM N.A.
RX MEDLINE=96090243; PubMed=7581362;
RA Harshtman K., Bell R., Rosenthal J., Katcher H., Miki Y.,
RA Swenson J., Gholami Z., Frye C., Ding W., Dayananth P.,
RA Edgington K., Norris F.H., Bristow P.K., Phelps R., Hattier T.,
RA Stone S., Shaffer D., Bayer S., Hussey C., Tian T., Lai M.,

```

```

RA Rostack P.R., Skolnick M.H., Shattuck-Eidens D., Kamb A.;
RT "Comparison of the positional cloning methods used to isolate the
RT BRCA1 gene.";
RL Hum. Mol. Genet. 4:1259-1266(1995).
CC -1- SIMILARITY: CONTAINS 1 ZF-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC or send an email to license@lsb-stb.ch).
CC -----
DR EMBL: X76952; CAAS4274.1;
DR EMBL: D30756; BAA06417.1;
DR EMBL: U25764; AAA93228.1;
DR GLYCOSITEDB: Q14596;
DR MIM: 166945;
DR InterPro: IPR000270; OPR.
DR InterPro: IPR000449; UBA.
DR InterPro: IPR000433; Znf_Z2.
DR Pfam: PF00564; OPR; 1.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF00569; Z2; 1.
DR SMART: SM00016; OPR; 1.
DR SMART: SM00291; Znf_Z2; 1.
DR PROSITE: PS01357; ZF_Z2_1;
DR PROSITE: PS01355; ZF_Z2_2; 1.
KW Antigen; Zinc-finger.
FT ZN_FING 211 257 Z2-TYPE.
FT DOMAIN 703 714 POLY-GLU.
FT CONFLICT 746 770 WGLCTALRSRHSQAMSEVTKASLGF -> LGDSMYSSALSG
FT CONFLICT 923 923 R -> H (IN REF. 2).
SQ SEQUENCE 966 AA; 107686 MW; 70A16F51E73D51A1 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 966;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SVPLVPC 9
Db 135 SVPLVPC 141

RESULT 11
TSPL_XENLA STANDARD; PRT; 1173 AA.
ID TSPL_XENLA
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desjardine D.W., Lawler J.;
RL Submitted (XX-1993) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND
CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,
CC LAMININ, TYPE V COLLAGEN AND INTEGRINS ALPHA-V/BETA-1, ALPHA-
CC V/BETA-3 AND ALPHA-IIb/BETA-3 (BY SIMILARITY).
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WFPC DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.

```

CC	-1- SIMILARITY:	CONTAINS 3 TSP TYPE-1 DOMAINS.
CC	-1- SIMILARITY:	CONTAINS 7 TSP TYPE-3 DOMAINS.
CC	-----	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/announce/ or send an email to license@isdb.sib.ch).	
CC	-----	
DR	EMBL:	L04278; ; NOT_ANNOTATED_CDS.
DR	HSSP:	P00740; 1IXA.
DR	InterPro:	IPR000561; EGF-like.
DR	InterPro:	IPR001861; EGF_Ca.
DR	InterPro:	IPR001791; Laminin_G.
DR	InterPro:	IPR000884; TSPI.
DR	InterPro:	IPR003129; TSPN.
DR	InterPro:	IPR001007; VMFC.
DR	InterPro:	IPR003367; tsp_3.
DR	pfam:	PF00008; EGF_2.
DR	pfam:	PF02210; TSPN; 1.
DR	pfam:	PEF00090; tsp_1; 3.
DR	pfam:	PF02412; tsp_3; 8.
DR	pfam:	PF00093; vmc; 1.
DR	SMART:	SM00181; EGF_2.
DR	SMART:	SM00001; EGF_like; 1.
DR	SMART:	SM00209; TSPI; 3.
DR	SMART:	SM00210; TSPN; 1.
DR	SMART:	SM00214; VMC; 1.
DR	PROSITE:	PS00022; EGF_1; FALSE_NEG.
DR	PROSITE:	PS01082; EGF_2; 1.
DR	PROSITE:	PS01086; TSPI; 3.
DR	PROSITE:	PS01208; VMFC; 1.
KW	Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;	
KM	EGF-like domain; Signal.	
FT	SIGNAL	1 22 POTENTIAL.
FT	CHAIN	23 1173 THROMBOSPONDIN 1.
FT	DOMAIN	23 235 HEPARIN-BINDING (POTENTIAL).
FT	DOMAIN	319 376 VMFC.
FT	DOMAIN	382 433 TSP TYPE-1 1.
FT	DOMAIN	438 494 TSP TYPE-1 2.
FT	DOMAIN	495 546 TSP TYPE-1 3.
FT	DOMAIN	550 590 EGF-LIKE 1.
FT	DOMAIN	591 648 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	649 693 EGF-LIKE 3.
FT	DOMAIN	726 761 TSP TYPE-3 1.
FT	DOMAIN	762 784 TSP TYPE-3 2.
FT	DOMAIN	785 820 TSP TYPE-3 3.
FT	DOMAIN	821 843 TSP TYPE-3 4.
FT	DOMAIN	844 881 TSP TYPE-3 5.
FT	DOMAIN	882 917 TSP TYPE-3 6.
FT	DOMAIN	918 953 TSP TYPE-3 7.
FT	DOMAIN	954 1173 C-TERMINAL.
FT	SITE	929 931 CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	554 565 BY SIMILARITY.
FT	DISULFID	559 575 BY SIMILARITY.
FT	DISULFID	578 589 BY SIMILARITY.
FT	DISULFID	595 621 BY SIMILARITY.
FT	DISULFID	602 610 BY SIMILARITY.
FT	DISULFID	623 647 BY SIMILARITY.
FT	DISULFID	653 666 BY SIMILARITY.
FT	DISULFID	660 679 BY SIMILARITY.
FT	DISULFID	681 692 BY SIMILARITY.
FT	CARBOHYD	155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	705 705 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	711 711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1070 1070 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	1173 AA; 130019 MW; A9F036D6516C0F24 CRC64;

Query Match	65.4%	Score 34	DB 1	Length 1173	
Best Local Similarity	62.5%	Pred. No. 1.3e+02			
Matches	5	Conservative	2	Mismatches	1
				Indels	0
				Gaps	0

QY 2 ISVPLVPC 9
: 11:11
Db 354 VSCPMP 361

RESULT 12
NU4M_ACACA STANDARD; PRT; 497 AA.
AC 037375;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
ND4 OR NAD4.
OS Acanthamoeba castellanii (Amoeba).
OC Mitochondrion.
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RSTR STRAIN=ATCC 30010 / NEFF;
RX MEDLINE=95147275; PubMed=7844823;
RA Burger G., Plante I., Loneragan K.W., Gray M.W.;
RT "The mitochondrial DNA of the amoeboid protozoon, Acanthamoeba
castellanii: complete sequence, gene content and genome
organization.";
RRL J. Mol. Biol. 245:522-537(1995).
RL -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; U12386; AAD11826.1; -
DR InterPro: IPR003918; NADHub_oxidrcdtsed4.
DR InterPro: IPR001750; Oxidored_q1.
DR Pfam: PF00361; Oxidored_q1: 1.
DR PRINTS: PR01437; N00XDRDTASE4.
KW Oxidoreductase; NAD; ubiquinone; Mitochondrion.
SQ SEQUENCE 497 AA; 57768 MW; 9C0B637E5C0F3B5 CRC64;

Query Match 63.5% Score 33; DB 1; Length 497;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CISVPLVP 8
: 1111
Db 230 CVKVPPLP 237

RESULT 13
CTPG_MYCTU STANDARD; PRT; 771 AA.
AC 010866;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable cation-transporting ATPase G (EC 3.6.3.-).
CTPG OR RV1992C OR MT2048 OR MTCY39.27.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;

[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean A., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Stalton J.E., Taylor K., Whitehead S., Barrett B.G., *et al.*
 RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES). SUBFAMILY IB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z74025; CA98411.1;
 DR EMBL: AE007056; AAK46325.1;
 DR TIGR: MT2048;
 DR TubercuList: RV1992c;
 DR InterPro: IPR001757; E1-E2_ATPase.
 DR InterPro: IPR001454; Hydrolase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATAPASE.
 DR PROSITE: PS00154; ATPASE_E1_E2; 1.
 DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KM Complete proteome.
 KW TRANSMEM 72
 FT TRANSMEM 163
 FT TRANSMEM 209
 FT TRANSMEM 330
 FT TRANSMEM 387
 FT TRANSMEM 411
 FT TRANSMEM 431
 FT TRANSMEM 457
 FT TRANSMEM 462
 FT MOD_RES 651
 FT METAL 655
 FT METAL 655
 SO SEQUENCE 771 AA; 79309 MW; 31281C919CC116CA CRC64;

Query Match
 Best Local Similarity 63.5%; Score 33; DB 1; Length 771;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1-CLVPLVP 8
 DB 377 CLVPLVP 384

RESULT 14
 ID DPOL_NEUCR
 AC P33537;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Probable DNA polymerase (EC 2.7.7.7).
 OS Neurospora crassa.
 OG Mitochondrion.
 OG Plasmid maranhar.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_Taxid=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AREY-1E;
 RX MEDLINE=93046810; PubMed=1423726;
 RA Court D.A., Bertrand H.;
 RT "Genetic organization and structural features of maranhar, a senescence-inducing linear mitochondrial plasmid of Neurospora crassa."
 RT Curr. Genet. 22:385-397(1992).
 RL -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + [DNA](N).
 CC -1- MISCELLANEOUS: THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. STRONG, TO DNA POLYMERASE OF OTHER FUNGAL AND PLANTS MITOCHONDRIAL PLASMIDS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X55361; CA93046.1;
 DR PIR: S26985; S26985.
 DR InterPro: IPR002064; DNA_POL_B.
 DR PRINTS: PR0106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Plasmid; Mitochondrion.
 SO SEQUENCE 1021 AA; 119075 MW; 54B9AC51CF5FE3A CRC64;

Query Match
 Best Local Similarity 50.0%; Score 33; DB 1; Length 1021;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISVPLVPC 9
 DB 670 INIPLVPC 677

RESULT 15
 ID PSTA_DICDI
 AC P11976;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE prestalk protein precursor.
 DE ECMB.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
 OX NCBI_Taxid=44689;
 RN [1]
 RP SEQUENCE FROM N.A.

Query Match
 Best Local Similarity 63.5%; Score 33; DB 1; Length 1021;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;


```

RC STRAIN-AX2;
RX MEDLINE=88015608; PubMed=3658700;
RA Ceccarelli A., McRobbie S.J., Jermyn K.A., Duffy K., Early A.,
RT Williams J.G.;
RT "Structural and functional characterization of a Dictyostelium gene
RT encoding a DIF inducible, prestalk-enriched mRNA sequence.";
RL Nucleic Acids Res. 15:7463-7476(1987).
RN [2]
RP SEQUENCE OF 1-17 FROM N.A.
RX MEDLINE=87257883; PubMed=3600646;
RA Ayres K., Neuman W., Rowekamp W.G., Chung S.;
RT "Developmental regulation of Dnase I-hypersensitive sites in
RT Dictyostelium discoideum.";
RL Mol. Cell. Biol. 7:1823-1829(1987).
CC -1- INDUCTION: BY THE PUTATIVE STALK-SPECIFIC MORPHOGEN DIF
CC (DIFFERENTIATION INDUCING FACTOR).
CC -1- SIMILARITY: THE REPEATED SEQUENCES ARE HIGHLY HOMOLOGOUS TO A
CC REPEAT OF THE PDD63 MRNA OF DICTYOSTELIUM DISCOIDEUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M16345; AAA33216.1; -.
DR PIR, A26838; A26838.
DR HSP, P00592; 3P2P.
DR DictyDb; DD03006; ecmb.
DR InterPro; IPR001673; S_mold_repeat.
DR Pfam; PF00526; S_mold_repeat; 42.
DR ProDom; PD006869; S_mold_repeat; 38.
KW Signal; Repeat.
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 1046 PRESTALK PROTEIN.
FT DOMAIN 31 1012 41 X 24 AA TANDEM REPEATS, CYS-RICH.
FT REPEAT 31 53 X-1.
FT REPEAT 54 77 X-2.
FT REPEAT 78 101 X-3.
FT REPEAT 102 125 X-1.
FT REPEAT 126 149 A-2.
FT REPEAT 150 173 B-1.
FT REPEAT 174 197 A-3.
FT REPEAT 198 221 A-4.
FT REPEAT 222 245 B-2.
FT REPEAT 246 269 A-5.
FT REPEAT 270 293 A-6.
FT REPEAT 294 317 B-3.
FT REPEAT 318 341 A-7.
FT REPEAT 342 365 A-8.
FT REPEAT 366 389 B-4.
FT REPEAT 390 413 A-9.
FT REPEAT 414 437 A-10.
FT REPEAT 438 461 B-5.
FT REPEAT 462 485 A-11.
FT REPEAT 486 509 A-12.
FT REPEAT 510 533 B-6.
FT REPEAT 534 557 A-13.
FT REPEAT 558 581 X-4.
FT REPEAT 582 605 A-14.
FT REPEAT 606 629 A-15.
FT REPEAT 630 653 A-16.
FT REPEAT 654 677 B-7.
FT REPEAT 678 701 A-17.
FT REPEAT 702 725 A-18.
FT REPEAT 726 749 B-8.
FT REPEAT 750 773 A-19.
FT REPEAT 774 797 A-20.
FT REPEAT 798 821 B-9.
FT REPEAT 822 845 A-21.
FT REPEAT 846 869 A-22.

```

```

FT REPEAT 870 893 A-23.
FT REPEAT 894 917 B-10.
FT REPEAT 918 941 A-24.
FT REPEAT 942 964 A-25.
FT REPEAT 965 988 A-26.
FT REPEAT 989 1012 A-27.
SQ SEQUENCE 1046 AA; 108823 MW; F9D6C04FD7D85818 CRC64;

Query Match      63.5%; Score 33; DB 1; Length 1046;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTSVPLVPC 9
DB 1028 C1HVPMTQC 1036

```

Search completed: June 17, 2002, 16:24:33
 Job time: 1562 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:23:45 ; Search time 172.85 Seconds
(without alignments)
9.008 Million cell updates/sec

Title: US-09-761-636a-14

Perfect score: 52
Sequence: 1 CISVPLVPC 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	78.8	342	2	Q53028
2	40	76.9	648	10	Q9MA88
3	39	75.0	297	5	P91378
4	39	75.0	347	16	Q9JY18
5	39	75.0	347	16	Q9JYU6
6	38	73.1	80	2	Q9Z1Y6
7	38	73.1	154	16	Q9PEZ3
8	38	73.1	198	2	Q30782
9	38	73.1	551	5	Q966C5
10	38	73.1	562	5	Q9VW94
11	38	73.1	1015	10	Q22243
12	38	73.1	3623	11	Q70244
13	37	71.2	287	11	Q9QY31
14	37	71.2	287	16	Q25921
15	37	71.2	407	5	Q93368
16	36	69.2	114	12	Q96736

17	36	69.2	185	16	Q99Y31	Q99Y31 streptococ
18	36	69.2	556	5	O15716	O15716 dictyostell
19	36	69.2	945	10	Q9M7A8	Q9M7A8 nicotiana t
20	36	69.2	1639	3	Q9P707	Q9P707 schizosacch
21	35	67.3	143	4	Q9KOC0	Q9KOC0 homo sapien
22	35	67.3	197	5	Q9BKRO	Q9BKRO caenorhabdi
23	35	67.3	249	4	Q9BPT7	Q9BPT7 homo sapien
24	35	67.3	265	4	Q9BPT4	Q9BPT4 homo sapien
25	35	67.3	305	16	Q9PEG0	Q9PEG0 xylella fas
26	35	67.3	309	4	Q9BX88	Q9BX88 homo sapien
27	35	67.3	388	5	Q96V3	Q96V3 strongyloid
28	35	67.3	496	10	Q01863	Q01863 phytophthor
29	35	67.3	551	11	Q88446	Q88446 ratius norv
30	35	67.3	556	4	O75751	O75751 homo sapien
31	35	67.3	557	4	Q96KH7	Q96KH7 homo sapien
32	35	67.3	558	11	Q9CX72	Q9CX72 mus musculu
33	35	67.3	565	4	Q9NVY3	Q9NVY3 homo sapien
34	35	67.3	569	13	O57339	O57339 xenopus lae
35	35	67.3	706	4	Q9BX90	Q9BX90 homo sapien
36	35	67.3	706	4	Q96SX2	Q96SX2 homo sapien
37	35	67.3	1108	2	P71485	P71485 mycobacteri
38	35	67.3	1210	5	O17241	O17241 caenorhabdi
39	35	67.3	1357	12	O89328	O89328 rice ragged
40	35	67.3	2240	2	Q9S129	Q9S129 streptomyce
41	34	65.4	158	10	Q9FMC3	Q9FMC3 arabidopsis
42	34	65.4	434	11	Q91V11	Q91V11 mus musculu
43	34	65.4	708	10	Q9LWP3	Q9LWP3 oryza sativ
44	34	65.4	715	13	Q98SP7	Q98SP7 brachydanio
45	34	65.4	789	5	Q9BL70	Q9BL70 caenorhabdi

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	342 AA.
Q53028	Q53028			
AC	Q53028			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	REDUCTASE.			
GN	AMOD.			
OS	Rhodococcus corallinus.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Corynebacterineae; Gordoniaceae; Gordonia.			
OX	NCBI_TaxID=36822;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-B-276;			
RA	Saeki H., Furuhashi K.;			
RT	"Cloning and characterization of a Nocardia corallina B-276 gene			
RT	cluster encoding alkene monooxygenase.";			
RL	J. Ferment. Biolog. 78:339-406(1994).			
CC	-I-CORFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).			
DR	EMBL: D37875; BAA07115.1; .			
DR	HSSP: P23486; IOFU.			
DR	InterPro: IPR000564; 2FE2S_ferredoxin.			
DR	InterPro: IPR001834; Cyt. B5_reductase.			
DR	InterPro: IPR001041; Ferredoxin.			
DR	InterPro: IPR001433; Oxidored_FAD.			
DR	InterPro: IPR001221; Phe_hydroxylase.			
DR	Pfam: PF00970; FAD_binding_6; 1.			
DR	Pfam: PF00111; fer2; 1.			
DR	Pfam: PF00175; NAD_binding_1.			
DR	PRINTS: PR00410; PHEHYDRXLA.			
DR	PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.			
KW	Iron-sulfur.			
SO	SEQUENCE 342 AA; 37282 MW; 4651F38D116A8049 CRC64;			

Query Match 78.8% Score 41; DB 2; Length 342;
Best Local Similarity 77.8% Pred. No. 6.4;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
 DB 75 CASVPLEPC 83

RESULT 2

O9MA88 PRELIMINARY; PRT; 648 AA.
 AC O9MA88;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 GN T12H1.28 PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Rongning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC009177; AAF27037.1; -
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR003575; Small_GTPase.
 DR SMART: SM00010; small_GTPase; 1.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
 SQ SEQUENCE 648 AA; 73066 MW; D22FD23B724D8B7B CRC64;

Query Match 76.9%; Score 40; DB 10; Length 648;
 Best Local Similarity 66.7%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
 DB 227 CDFPLVPC 235

RESULT 3

P91378 PRELIMINARY; PRT; 297 AA.
 AC P91378;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 GN HYPOTHETICAL 33.8 KDA PROTEIN.
 GN K12D9.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peioderinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA GRAVES T.;
 RT "The sequence of C. elegans cosmid K12D9.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U80030; AAC24161.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 297 AA; 33823 MW; C300EDA72169ADB8D CRC64;

Query Match 75.0%; Score 39; DB 5; Length 297;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
 DB 169 C1SVPLVPC 177

RESULT 4

O9JY18 PRELIMINARY; PRT; 347 AA.
 AC O9JY18;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE.
 GN NMB1787.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NC NCB1_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=2015755; PubMed=10710307;
 RA Tettein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yatchew J.,
 RA Gill J., Scarlato V., Maignani V., Piazza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002528; AAF42126.1; -
 DR TIGR; NMB1787; -
 DR InterPro: IPR000706; AGPR_act_site.
 DR InterPro: IPR000534; Semialdh_dh.
 DR Pfam; PF01118; Semialdh_dh; 1.
 DR Pfam; PF02774; Semialdh_dh; 1.
 DR ProDom; PD003765; AGPR_act_site; 1.
 KW Complete proteome.
 SQ SEQUENCE 347 AA; 37213 MW; A927C570DE19ECD1 CRC64;

Query Match 75.0%; Score 39; DB 16; Length 347;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 8
 DB 156 C1SVPLVPC 163

RESULT 5
 O9JY06 PRELIMINARY; PRT; 347 AA.
 AC O9JY06;

DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PUTATIVE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE
 DE (EC 1.2.1.38).
 GN ARGC OR NMA0676.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NCBI_TaxID=56599;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / SEROTYPE 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churruarin C.,
 Ralee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Mould S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrell B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491".
 RL Nature 404:502-506(2000).
 DR EMBL: AL162753; CAB83963.1;
 DR InterPro: IPR000706; AGPR_act_site.
 DR InterPro: IPR000534; Semialdhc.
 DR Pfam: PF01118; Semialdehyde_dhc. 1.
 DR Pfam: PF02774; Semialdehyde_dhc. 1.
 DR ProDom: PD003765; AGPR_act_site; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 347 AA; 37227 MW; E11E8A19A2B9246 CRC64;

Query Match 75.0%; Score 39; DB 16; Length 347;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 C1SVPLVP 8
 Db 156 C1SVPLVP 163
 RESULT 6
 O92LY6 PRELIMINARY; PRT; 80 AA.
 AC O92LY6:
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DE HYPOTHEICAL 9.2 KDA PROTEIN.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N40;
 RA Feng S., Chen G., Barthold S.;
 RT "G5 clone from Borrelia burgdorferi.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF005036; AAD01230.1;
 KW Hypothetical protein.
 SQ SEQUENCE 80 AA; 9171 MW; C61A1E359DEDBEA3 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 80;
 Best Local Similarity 87.5%; Pred. No. 5.6;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 C1SVPLVP 8
 Db 40 C1SVPLVP 47
 RESULT 7

O9PFZ3
 ID O9PFZ3 PRELIMINARY; PRT; 154 AA.
 AC O9PFZ3:
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE PHAGE-RELATED ENDOLYSIN.
 DE XP0513.
 GN Xylella.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 OX [1]
 NCBI_TaxID=2371;
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaerenga R., Alves L.M.C., Arya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carriro D.M., Carer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miraca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
 RA Quagga R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa".
 RL Nature 406:151-159(2000).
 DR EMBL: AE003900; AAF83323.1;
 DR InterPro: IPR002196; Phage_lysozyme.
 DR Pfam: PF00959; Phage_lysozyme; 1.
 KW Complete proteome.
 SQ SEQUENCE 154 AA; 16793 MW; 20A5570A19D715FD CRC64;

Query Match 73.1%; Score 38; DB 16; Length 154;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 C1SVPLVP 8
 Db 78 C1SVPLVP 85

RESULT 8
 O30782 PRELIMINARY; PRT; 198 AA.
 ID O30782:
 AC O30782:
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DE INCLUSION MEMBRANE PROTEIN B.
 DE INCB.
 GN Chlamydomophila caviae.
 OS Chlamydiales; Chlamydiaceae; Chlamydomophila.
 OX NCBI_TaxID=83557;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-GPIC;
 RX MEDLINE-98326820; PubMed-9663687;
 RA Banantline J.P., Rokey D.D., Hackstadt T.;
 RT "Random genes of Chlamydia psittaci that encode proteins localized to
 the inclusion membrane.";
 RL Mol. Microbiol. 28:1017-1026(1998).
 DR EMBL: AF017105; AAC46378.1;
 SQ SEQUENCE 198 AA; 20917 MW; 455E5EA09444FE29 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 198;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 CISVPLVPC 9
 |||:|:|
 Db 157 CISLPIISC 165

RESULT 9

ID O966C5 PRELIMINARY; PRT; 551 AA.

AC O966C5;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHEICAL PROTEIN Y24D9A.2.
 GN Y24D9A.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_Taxid=6239;

RA None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Cordes M., Wohlmann P.;
 RT "The sequence of C. elegans cosmid Y24D9A.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006729; AAC84599.1;
 SQ SEQUENCE 551 AA; 63939 MW; 73BAD1C7DF1EB56F CRC64;

Query Match 73.1%; Score 38; DB 5; Length 551;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLVPC 9
 |||:|:|
 Db 260 CSSKPIVPC 268

RESULT 10

ID O9VN94 PRELIMINARY; PRT; 562 AA.

AC O9VN94;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)

DE CG1113 PROTEIN.
 GN CG1113.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mills G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencs P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Geldart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Jimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong E., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003603; AAF52054.1;
 DR FlyBase: FBgn0037304; CG1113.
 SQ SEQUENCE 562 AA; 63406 MW; AC09F5B6F5DE2F970 CRC64;

Query Match 73.1%; Score 38; DB 5; Length 562;
 Best Local Similarity 55.6%; Pred. No. 38;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CISVPLVPC 9
 |||:|:|
 Db 9 CVKVPIDPC 17

RESULT 11

ID O22243 PRELIMINARY; PRT; 1015 AA.

AC O22243;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-AUG-1999 (Tremblrel. 11, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE ATP-DEPENDENT RNA HELICASE A.
 GN ATG47680.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umeyama L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A.,
 RA Salberger S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 [12]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005309; AAC63624.1; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000571; Zf-CCH.
 DR Pfam: PF00371; Helicase_C; 1.
 DR Pfam: PF00642; Zf-CCH; 2.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR SMART: SM00356; ZNF_C3H1; 2.
 DR ATP-binding; Helicase.
 KW
 SQ SEQUENCE 1015 AA; 115084 MW; DIC342B38C561C9 CRC64;

Query Match 73.1%; Score 38; DB 10; Length 1015;
 Best Local Similarity 75.0%; Pred. No. 66;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 C1SVPLVP 8
 1:111 11
 DB 676 C1SVPLVP 683

RESULT 12
 ID 070244 PRELIMINARY; PRT; 3623 AA.
 AC 070244;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INTRINSIC FACTOR-B12 RECEPTOR PRECURSOR.
 GN CUBILIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98148073; PubMed=9478979;
 RA Moestrup S.K., Kozyraki R., Kristiansen M., Kayen J.H.,
 RA Rasmussen H.H., Brault D., Pontillon F., Goda F.O., Christensen E.I.,
 RA Hammond T.G., Verroust P.J.;
 RT "The intrinsic factor-vitamin B12 receptor and target of teratogenic
 RT antibodies is a megalin-binding peripheral membrane protein with
 RT homology to developmental proteins.";
 RL J. Biol. Chem. 273:5235-5242(1998).
 CC -1- SIMILARITY: CONTAINS 26 CUB DOMAINS.
 DR EMBL: AF022247; AAC71661.1; -;
 DR HSSP: P00740; 1EDM.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000859; CUB.

DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00431; CUB; 27.
 DR Pfam: PF00008; EGF; 7.
 DR SMART: SM00042; CUB; 26.
 DR SMART: SM00179; EGF_CA; 4.
 DR SMART: SM00001; EGF-like; 4.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS01180; CUB; 27.
 DR PROSITE: PS00022; EGF_1; UNKNOWN; 4.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 4.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
 KW Receptor; Repeat; signal.
 FT SIGNAL 1 20
 FT CHAIN 21 3623
 SQ SEQUENCE 3623 AA; 398981 MW; 39FB792AC6545240 CRC64;

Query Match 73.1%; Score 38; DB 11; Length 3623;
 Best Local Similarity 66.7%; Pred. No. 2,2e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 C1SVPLVPC 9
 1:111 11
 DB 313 C1SVPLVPC 321

RESULT 13
 ID 090Y31 PRELIMINARY; PRT; 287 AA.
 AC 090Y31;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SNAIL-RELATED ZINC FINGER PROTEIN SMUC.
 GN ZFP293 OR SMUC.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6;
 RX MEDLINE=20075155; PubMed=10606664;
 RA Kataoka H., Murayama T., Tokode M., Mori S., Sano H., Ozaki H.,
 RA Yokota Y., Nishikawa S.I., Kita T.;
 RT "A novel Snail-related transcription factor Smuc regulates basic
 RT helix-loop-helix transcription factor activities via specific E-box
 RT motifs.";
 RL Nucleic Acids Res. 28:626-633(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AF133714; AAF22956.1; -;
 DR HSSP: P15822; 1BBO.
 DR MGD: MGI:1353563; Zfp293.
 DR InterPro: IPR002355; Multicn_oxidase2.
 DR InterPro: IPR000822; Znf-C2H2.
 DR Pfam: PF00096; Zf-C2H2; 5.
 DR PRINTS: PR00048; ZINCINGER.
 DR SMART: SM00355; ZNF_C2H2; 5.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE: PS00577; ZINC_FINGER_C2H2_2; 5.
 KW DNA-binding; Metal-binding; Nucleic acid protein; Zinc-finger.
 SQ SEQUENCE 287 AA; 31636 MW; 02E42A6F8BAFD48 CRC64;

Query Match 71.2%; Score 37; DB 11; Length 287;
 Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 C1SVPLVP 8

DB 67 CINSPLP 74

RESULT 14

025921 PRELIMINARY: PRT: 287 AA.
 AC 025921;
 DT 01-JAN-1998 (TRENBLREL. 05, Created)
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE TYPE IIS RESTRICTION ENZYME M2 PROTEIN (MOD).
 GN Hpi368.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khaliq H.G., Glodek A.,
 RA McEneney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori."
 RL Nature 388:539-547(1997).
 DR EMBL: AE000637; AAD08412.1; -.
 DR TIGR: Hpi368; -.
 DR InterPro: IPR002295; D2IN6.mtfase.
 DR InterPro: IPR001091; N6_M4_mase.
 DR InterPro: IPR002941; N6_M4_mase.
 DR Pfam: PF01555; N6_M4_mase; 1.
 DR PRINTS: PR00506; D2IN6MTFRASE.
 DR PRINTS: PR00508; S2IN4MTFRASE.
 DR PROSITE: PS00093; N4_MTASE; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 287 AA; 33516 MW; 912B967A6ADA0653 CRC64;

Query Match 71.2%; Score 37; DB 16; Length 287;

Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CINSPLP 8
 DB 86 CINSPLP 93

RESULT 15

093368 PRELIMINARY: PRT: 407 AA.
 AC 093368;
 DT 01-FEB-1997 (TRENBLREL. 02, Created)
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE C41G7.4 PROTEIN.
 GN C41G7.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peioderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Steward C.A.;
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;

RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
 DR EMBL: Z81048; CAB02841.1; -.
 DR InterPro: IPR001214; SET.
 DR Pfam: PF00856; SET; 1.
 DR SMART: SM00317; SET; 1.
 DR PROSITE: PS50280; SET; 1.
 SQ SEQUENCE 407 AA; 46221 MW; B275994433254520 CRC64;

Query Match

Best Local Similarity 71.2%; Score 37; DB 5; Length 407;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CINSPLP 9
 DB 113 CINSPLP 121

Search completed: June 17, 2002, 16:23:47
 Job time: 1621 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2002, 16:02:17 ; Search time 227.35 Seconds
(without alignments)
4.397 Million cell updates/sec

Title: US-09-761-636a-14

Perfect score: 52
Sequence: 1 C1SWPLVPC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	9	22	AAU04533
2	41.5	79.8	10	22	AAU04532
3	41	78.8	11	22	AAU04526
4	41	78.8	342	16	AA065216
5	41	78.8	342	17	AA081472
6	40	76.9	11	22	AAU04545
7	39	75.0	11	22	AAU04542
8	39	75.0	101	22	AA06501
9	39	75.0	102	22	AB022400
10	38	73.1	44	20	AA027025
11	38	73.1	198	20	AA032174

12	38	73.1	562	22	ABB57779	Drosophila melanog
13	38	73.1	3623	20	AA027020	Rat cubilin proteol
14	37	71.2	64	21	AA004558	Human secreted pro
15	37	71.2	65	20	AA012860	Human 5' EST secre
16	36	69.2	11	22	AAU04544	VEGF based monocyc
17	35	67.3	56	20	AA085729	Novel protein (C10
18	35	67.3	120	22	AA093613	Human polypeptide,
19	35	67.3	200	22	AB036676	Novel human diagno
20	35	67.3	249	19	AA070900	Human secreted pro
21	35	67.3	249	22	AAU09180	Human PRO1884 poly
22	35	67.3	249	22	AAU12226	Human PRO1884 poly
23	35	67.3	251	21	AA041885	Human ORFX ORF1649
24	35	67.3	287	22	AA025559	Human G Protein-Co
25	35	67.3	309	21	AA032332	Eucalyptus grandis
26	35	67.3	313	22	AAU25556	Human G Protein-Co
27	35	67.3	353	22	AAU10068	Chemokine receptor
28	35	67.3	370	21	AA032774	Eucalyptus grandis
29	35	67.3	372	22	AAU10067	Chemokine receptor
30	35	67.3	482	22	AB031666	Novel human diagno
31	35	67.3	521	22	AB074332	Novel human diagno
32	35	67.3	532	22	AB014722	Novel human diagno
33	35	67.3	565	22	AB011289	Novel human diagno
34	35	67.3	565	22	AB011291	Novel human diagno
35	35	67.3	565	22	AA092822	Human protein sequ
36	35	67.3	706	22	AA094174	Human protein sequ
37	35	67.3	1512	22	AA078732	Human protein SEQ
38	35	67.3	1512	22	AAU04349	Mammalian toxicolo
39	35	67.3	1550	22	AA079716	Human protein SEQ
40	34	65.4	11	22	AAU04543	VEGF based monocyc
41	34	65.4	61	22	AA090405	Human immune/haema
42	34	65.4	67	22	AAU40341	Propionibacterium
43	34	65.4	109	22	AB012918	Novel human diagno
44	34	65.4	118	22	AAU3981	Propionibacterium
45	34	65.4	161	22	AB04230	Peptide #11736 enc

ALIGNMENTS

RESULT 1	
AAU04533	standard; Peptide; 9 AA.
ID	AAU04533
AC	AAU04533;
XX	26-SEP-2001 (first entry)
DT	VEGF based monocyclic peptide 11.
DE	Human: VEGF; vascular endothelial growth factor; angiogenesis;
XX	neovascularisation; lymphangiogenesis; psoriasis; tumour;
KW	diabetes induced neovascular sequelae; rheumatoid arthritis;
KW	diabetic retinopathy; chronic inflammation; cyclic.
OS	Synthetic.
XX	
FT	Key
FT	Disulfide bond 1..9
FT	/note= "This bond cyclises the peptide"
FN	WO200152875-A1.
XX	
PD	26-JUL-2001.
XX	
PF	18-JAN-2001; 2001WO-US01533.
XX	
PR	18-JAN-2000; 2000US-0176293.
XX	
PR	16-MAY-2000; 2000US-0204590.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	
XX	Achen MG, Hughes RA, Stacker S, Cendron A;
PI	
XX	

DR WPI: 2001-442248/47.
 XX
 PF Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Claim 49; Page 32; 102pp; English.
 XX
 CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 XX
 SO Sequence 9 AA:
 Query Match 100.0%; Score 52; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTSVPLVPC 9
 |||||
 DB 1 CTSVPLVPC 9
 |||||
 RESULT 2
 AA004532
 ID AA004532 standard; Peptide: 10 AA.
 XX
 AC AA004532;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based monocyclic peptide 10.
 XX
 KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.
 XX
 OS Synthetic.
 XX
 AC Key
 XX Location/Qualifiers
 FT Disulfide-bond 1..10
 XX /note- "this bond cyclises the peptide"
 XX
 PN W0200152875-A1.
 XX
 PD 26-JUL-2001.

XX
 PF 18-JAN-2001; 2001WO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 XX
 DR WPI: 2001-442248/47.
 XX
 XX Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Claim 49; Page 32; 102pp; English.
 XX
 CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 XX
 SO Sequence 10 AA:
 Query Match 79.8%; Score 41.5; DB 22; Length 10;
 Best Local Similarity 90.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CTSVPL-VPC 9
 |||||
 DB 1 CTSVPLVPC 10
 |||||
 RESULT 3
 AA004526
 ID AA004526 standard; Peptide: 11 AA.
 XX
 AC AA004526;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based monocyclic peptide 3.
 XX
 KW Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KW neovascularisation; lymphangiogenesis; psoriasis; tumour;
 KW diabetes induced neovascular sequelae; rheumatoid arthritis;
 KW diabetic retinopathy; chronic inflammation; cyclic.

XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Disulfide-bond 1..11
FT /note="This bond cyclises the peptide"
XX PN WO200152875-A1.
XX PD 26-JUL-2001.
XX PE 18-JAN-2001; 2001WO-US01533.
XX PR 18-JAN-2000; 2000US-0176293.
XX PR 16-MAY-2000; 2000US-0204590.
XX PA (LUDWIG INST CANCER RES.
PI Achen MG, Hughes RA, Stacker S, Cendron A;
PI WPI: 2001-442248/47.
XX DR
XX PT Novel monomeric monocyclic peptide, used to interfere with
PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
PT loop fragment from an exposed loop of a growth factor protein by
PT oxidizing the cysteine residues -
XX PS
XX PS Claim 49; Page 32; 102pp; English.
XX CC The sequence represents a monomeric monocyclic peptide of the invention,
XX CC whose 3-dimensional structure is modelled on the expose loop of human
XX CC VEGFD (vascular endothelial growth factor). The invention relates to a
XX CC method of producing a monomeric monocyclic peptide by a measuring
XX CC beta-beta carbon separation distances on opposite antiparallel strands of
XX CC a peptide loop fragment from an exposed loop of a growth factor protein
XX CC and cyclising the peptide by oxidising the cysteine residues. The
XX CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
XX CC monocyclic peptides) and a cyclic peptide with at least one amino acid
XX CC deleted prior to cyclisation are used to interfere with angiogenesis,
XX CC neovascularisation or lymphangiogenesis in a mammal with a condition
XX CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
XX CC The condition is diabetic retinopathy, psoriasis, arthropathy,
XX CC hemangioma, vascularised malignant or benign tumour, post-recovery
XX CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
XX CC trauma, substance-induced neovascularisation of the liver, excessive
XX CC hormone-related angiogenic dysfunction, diabetes induced neovascular
XX CC sequelae, hypertension induced neovascular sequelae, or chronic liver
XX CC infection. The peptides are also used to modulate vascular permeability
XX CC in a mammal (the mammal has a condition characterised by fluid
XX CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
XX CC or brain. The peptides are used to image blood vessels and lymphatic
XX CC vasculature. The monomeric and bicyclic peptides are used to interfere
XX CC with at least one biological activity induced by VEGF, VEGF-C or -D and
XX CC are also used in combination with an anti-inflammatory agent, to treat a
XX CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
XX CC diabetic retinopathy.
XX SQ Sequence .11 AA;

Query Match 78.8%; Score 41; DB 22; Length 11;
Best Local Similarity 81.8%; Pred. NO. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 CTSVPL--VPC 9
DB 1 CTSVPLSVPC 11

RESULT 4
AAR66216
ID AAR66216 standard; Protein: 342 AA.
XX

AC AAR66216;
XX
XX 04-AUG-1995 (first entry)
XX DE Nocardia corallina reductase (encoded by amod).
XX KM alkene monooxygenase; reductase; epoxidation; amod.
XX OS Nocardia corallina.
XX PN JP06292571-A.
XX PD 21-OCT-1994.
XX PE 06-APR-1993; 93JP-0105171.
XX PR 06-APR-1993; 93JP-0105171.
XX PA (NIHA) JAPAN ENERGY CORP.
XX DR WPI: 1995-009069/02.
XX DR N-PSDB; AAQ79569.
XX PT Alkene monooxygenase and corresp. gene - useful for the
XX PT epoxidation of an alkene
XX PS Claim 2; Page 4-5; 30pp; Japanese.
XX CC E.coli transformed with the DNA sequence AAQ79569 are able to
XX CC catalyse the epoxidation of alkenes. The DNA is derived from
XX CC Nocardia corallina and comprises 4 open reading frames. ORFs amoa
XX CC and amoc encode subunits 1 and 2 of the alkene monooxygenase enzyme
XX CC from NADH coenzyme to a monooxygenase.
XX SQ Sequence 342 AA;

Query Match 78.8%; Score 41; DB 16; Length 342;
Best Local Similarity 77.8%; Pred. NO. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTSVPLVPC 9
DB 75 casvplpc 83

RESULT 5
AAR81472
ID AAR81472 standard; Protein: 342 AA.
XX AAR81472:
XX 07-AUG-1996 (first entry)
XX DE Nocardia corallina alkene mono-oxygenase gene product, amod.
XX KM Alkene mono-oxygenase; indole; indigo production; biosynthesis;
XX KM microbial oxidation; dye.
XX OS Nocardia corallina B-276.
XX PN JP08023988-A.
XX PD 30-JAN-1996.
XX PE 08-JUL-1994; 94JP-0179688.
XX PR 08-JUL-1994; 94JP-0179688.
XX PA (NIHA) JAPAN ENERGY CORP.
XX DR WPI: 1996-133426/14.

DR N-PSDB: AAT17418.
 XX
 XX Prepn. of indigo by a microbiological method - by culturing a
 PT microbe having alkene monooxygenase activity to oxidise indole to
 PT indigo
 XX
 PS Claim 5; Page 8-9; 11pp: Japanese.
 XX
 CC AAR81469-R81472 are protein products of the Nocardia corallina strain
 CC B-276 alkene mono-oxygenase gene. The gene encodes 4 protein products
 CC amOA, amOB, amOC and amOD derived from the 3 different reading frames
 CC of the operon. The gene is useful for the production of indigo via
 CC oxidation of indole. Nocardia corallina can be cultured in a medium
 CC contg. indole and will readily oxidise the indole yielding indigo into
 CC the culture medium. E. coli may also be transformed with the alkene
 CC mono-oxygenase gene and used as above to efficiently produce indigo
 CC by microbial oxidation.
 CC
 XX
 SQ Sequence 342 AA;
 Query Match 78.8%; Score 41; DB 17; Length 342;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CTSVPLVPC 9
 1 | 1111 | 11
 DB 75 casvpltpc 83
 RESULT 6
 AAU04545
 ID AAU04545 standard; Peptide: 11 AA.
 AC AAU04545;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based monocyclic peptide 23.
 XX
 KM Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KM neovascularisation; lymphangiogenesis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..11
 FT /note= "This bond cyclises the peptide"
 XX
 PN WO200152875-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001MO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 DR WPI; 2001-442248/47.
 XX
 PT Novel monomeric monocyclic peptide, used to interfere with
 PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide
 PT loop fragment from an exposed loop of a growth factor protein by
 PT oxidizing the cysteine residues -
 XX
 PS Example 25; Page 47; 102pp; English.
 XX

CC The sequence represents a monomeric monocyclic peptide of the invention,
 CC whose 3-dimensional structure is modelled on the expose loop of human
 CC VEGF (vascular endothelial growth factor). The invention relates to a
 CC method of producing a monomeric monocyclic peptide by a measuring
 CC beta-beta carbon separation distances on opposite antiparallel strands of
 CC a peptide loop fragment from an exposed loop of a growth factor protein
 CC and cyclising the peptide by oxidising the cysteine residues. The
 CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked
 CC monocyclic peptides) and a cyclic peptide with at least one amino acid
 CC deleted prior to cyclisation are used to interfere with angiogenesis,
 CC neovascularisation or lymphangiogenesis in a mammal with a condition
 CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.
 CC The condition is diabetic retinopathy, psoriasis, arthropathy,
 CC hemangioma, vascularised malignant or benign tumour, post-recovery
 CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold
 CC trauma, substance-induced neovascularisation of the liver, excessive
 CC hormone-related angiogenic dysfunction, diabetes induced neovascular
 CC sequelae, hypertension induced neovascular sequelae, or chronic liver
 CC infection. The peptides are also used to modulate vascular permeability
 CC in a mammal (the mammal has a condition characterised by fluid
 CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,
 CC or brain. The peptides are used to image blood vessels and lymphatic
 CC vasculature. The monomeric and bicyclic peptides are used to interfere
 CC with at least one biological activity induced by VEGF, VEGF-C or -D and
 CC are also used in combination with an anti-inflammatory agent, to treat a
 CC chronic inflammation, especially rheumatoid arthritis, psoriasis and
 CC diabetic retinopathy.
 CC
 XX
 SQ Sequence 11 AA;
 Query Match 76.9%; Score 40; DB 22; Length 11;
 Best Local Similarity 72.7%; Pred. No. 2;
 Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
 QY 1 CTSVPL--VPC 9
 1 | 1111 | 11
 DB 1 casvpltpc 11
 RESULT 7
 AAU04542
 ID AAU04542 standard; Peptide: 11 AA.
 AC AAU04542;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE VEGF based monocyclic peptide 20.
 XX
 KM Human; VEGF; vascular endothelial growth factor; angiogenesis;
 KM neovascularisation; lymphangiogenesis; tumour;
 KM diabetes induced neovascular sequelae; rheumatoid arthritis;
 KM diabetic retinopathy; chronic inflammation; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..11
 FT /note= "This bond cyclises the peptide"
 XX
 PN WO200152875-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001MO-US01533.
 XX
 PR 18-JAN-2000; 2000US-0176293.
 PR 16-MAY-2000; 2000US-0204590.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Achen MG, Hughes RA, Stacker S, Cendron A;
 XX

DR WP1: 2001-44a2248/47.

XX Novel monomeric monocyclic peptide, used to interfere with

PT angiogenesis, or lymphangiogenesis, is produced by cyclising a peptide

PT loop fragment from an exposed loop of a growth factor protein by

PT oxidizing the cysteine residues -

XX

PS Example 25; Page 47; 102pp: English.

XX

CC The sequence represents a monomeric monocyclic peptide of the invention,

CC whose 3-dimensional structure is modelled on the expose loop of human

CC VEGF₁₂₁ (vascular endothelial growth factor). The invention relates to a

CC method of producing a monomeric monocyclic peptide by a measuring

CC beta-beta carbon separation distances on opposite antiparallel strands of

CC a peptide loop fragment from an exposed loop of a growth factor protein

CC and cyclising the peptide by oxidising the cysteine residues. The

CC monocyclic peptides, dimeric bicyclic peptides (comprising 2 linked

CC monocyclic peptides) and a cyclic peptide with at least one amino acid

CC deleted prior to cyclisation are used to interfere with angiogenesis,

CC neovascularisation or lymphangiogenesis in a mammal with a condition

CC characterised by angiogenesis, neovascularisation or lymphangiogenesis.

CC The condition is diabetic retinopathy, psoriasis, arthropathy,

CC hemangioma, vascularised malignant or benign tumour, post-recovery

CC cerebrovascular accident, post-angioplasty restenosis, head, heat or cold

CC trauma, substance-induced neovascularisation of the liver, excessive

CC hormone-related angiogenic dysfunction, diabetes induced neovascular

CC sequelae, hypertension induced neovascular sequelae, or chronic liver

CC infection. The peptides are also used to modulate vascular permeability

CC in a mammal (the mammal has a condition characterised by fluid

CC accumulation in peripheral limbs or in lungs, peritoneal cavity, pleura,

CC or brain. The peptides are used to image blood vessels and lymphatic

CC vasculature. The monomeric and bicyclic peptides are used to interfere

CC with at least one biological activity induced by VEGF, VEGF-C or -D and

CC are also used in combination with an anti-inflammatory agent, to treat a

CC chronic inflammation, especially rheumatoid arthritis, psoriasis and

CC diabetic retinopathy.

XX

SO Sequence 11 AA:

75.0%; Score 39; DB 22; Length 11;

Query Match Best Local Similarity 72.7%; Pred. No. 2.9;

Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 C1SVPL--VPC 9

I:|||||

Db 1 clsvplstvc 11

RESULT 8

AA06501

ID AA06501 standard; Protein; 101 AA.

XX

AA06501;

XX

05-OCT-2001 (first entry)

DE Human foetal protein, SEQ ID NO: 232.

XX

Human: foetal protein; cytostatic; immunosuppressive; immunostimulant;

KW neotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;

KW gene therapy; antisense therapy; cancer; immune disorder;

KW growth disorder; osteoporosis; thrombolytic disorder;

KW nervous system disorder; Inflammation.

XX

XX Homo sapiens.

PN WO200155339-A2.

XX

02-AUG-2001.

XX

25-JAN-2001; 2001WO-US02723.

```

PR 25-JAN-2000; 2000US-0491404.
PR 15-SEP-2000; 2000US-0663870.
PR 06-NOV-2000; 2000US-0707351.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Yeung G, Ford JE, Boyle BJ, Atterburn MC, Drmanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Werman T;
XX
XX WPI: 2001-465571/50.
DR N-PSDB; AAH94176.
XX
PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation -
XX
XX Claim 10; Page 264; 715pp; English.
XX
XX The invention relates to novel foetal polypeptides encoded by
XX CC polynucleotides comprising one of 477 sequences fully defined in the
XX CC specification. The foetal polynucleotides and polypeptides are
XX CC useful in the treatment and diagnosis of diseases such as cancers,
XX CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
XX CC disorders, nervous system disorders and inflammation. The present
XX CC sequence is a polypeptide encoded by a cDNA assembled using
XX CC an expressed sequence tag (EST) found to be expressed in human
XX CC foetal tissue cDNA libraries.
XX
SQ Sequence 101 AA;
XX
Query Match 75.0%; Score 39; DB 22; Length 101;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CISVPLVP 8
| | | | | |
Db 2 cisvpltp 9
| | | | | |
RESULT 9
ABG22400
ID ABG22400 standard; Protein; 102 AA.
XX
AC ABG22400;
XX
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #22391.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
WPI: 2001-639362/73.
DR N-PSDB; AAS86587.
XX
```

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 52759; 103bp; English.

XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 102 AA;

Query Match 75.0%; Score 39; DB 22; Length 102;

Best Local Similarity 87.5%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTSVPLVP 8
| | | | |
Db 24 cfsvplvp 31

RESULT 10

AAV27025
ID AAV27025 standard; Protein: 44 AA.

AC AAV27025;

DT 08-OCT-1999 (first entry)

XX Amino acid sequence of rat cubilin EGF5 repeat.

XX Cubilin; epithelial glycoprotein receptor; vitamin B12; endocytosis;
KW toxicity; kidney; spleen; brain; liver; heart; thyroid; abortion; rat;
KW proteinuria; fetal malformation; fetal development; kidney damage; EGF.
XX
OS Rattus sp.

PN WO9937757-A1.

PD 29-JUL-1999.

PF 21-JAN-1999; 99WO-US01259.

PR 22-JAN-1998; 98US-0072197.

PA (INRM) INST NAT SANTE & RECH MEDICALE.
PA (TULA) TULANE EDUCATIONAL FUND.

PI Hammond TG, Verroust PJ;

WP: 1999-479045/40.

PT New DNA encoding cubilin, used for treating toxicity, particularly

PT nephrotoxicity, and as marker of kidney damage

XX Example 18; Fig 6A; 135bp; English.

XX
XX The invention relates to a rat cubilin protein. Cubilin is a 11gand-
CC binding, epithelial glycoprotein receptor that facilitates uptake of
CC intrinsic factor/vitamin B12 complexes in intestines and kidney. It is
CC also involved in endocytosis and trafficking of light immunoglobulin
CC chains in renal proximal tubule cells. Host cells containing a vector
CC comprising the rat cubilin DNA sequence can be used for the recombinant
CC expression of the protein. Cubilin, or its fragments, are used to treat
CC or reduce toxicity, particularly in kidneys, spleen, brain, liver, heart
CC and thyroid. Cubilin mutations may also be implicated in idiopathic
CC proteinuria, fetal malformation, poor fetal development and spontaneous
CC abortions. Cubilin may also be used to raise specific antibodies, used
CC for its detection, or clones that express it, in standard immunoassays.
CC Fragments of cubilin DNA can also be used to detect cubilin mRNA in cell
CC and tissues, by hybridization. Abnormal levels of cubilin in the urine
CC are indicative of kidney damage. Sequences AAV27021-35 represent EGF
CC repeat sequences of rat cubilin and from homologous regions of other
CC proteins.

SQ Sequence 44 AA;

Query Match 73.1%; Score 38; DB 20; Length 44;

Best Local Similarity 66.7%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CTSVPLVPC 9
| | | | |
Db 12 csqsvplvp 20

RESULT 11

AAV32174
ID AAV32174 standard; Protein: 198 AA.

AC AAV32174;

DT 01-FEB-2000 (first entry)

XX Chlamydia psittaci infection-specific protein IncB.

XX IncB; infection; vaccine; therapy; diagnosis.

XX Chlamydia psittaci.

PN WO9953948-A1.

PD 28-OCT-1999.

PF 20-APR-1999; 99WO-US08744.

PR 20-APR-1998; 98US-0082438.

PR 21-APR-1998; 98US-0082588.

PR 22-MAY-1998; 98US-0086450.

PA (UYOR-) UNIV OREGON STATE.

PI Rockey DD, Bannantine JP;

WP: 1999-633904/54.

DR N-PSDB; AA234588.

PT Novel bacterial infection specific proteins for treating and diagnosing
PT chlamydial infections

PS Claim 1; Page 41-42; 56bp; English.

CC This sequence represents novel infection-specific protein IncB of
CC Chlamydia psittaci strain GPIC. IncB is found in the inclusion
CC membrane of infected cells. It is associated primarily with the

vegetative reticulate body form of Chlamydia rather than with the
refractive elementary body form. The invention includes: a vaccine
directed against the reticulate body form of Chlamydia comprising 1
or more infection-specific proteins (see AAY32170-78), including
IncA, IncB and IncC; methods of using and producing such a vaccine;
methods for detection of infection-specific antibodies or antigens
in a biological specimen; and a method of using therapeutic agents
specifically directed against infection-specific peptides, or the
genes that code for such peptides, to treat chlamydial infection.

Sequence 198 AA:

Query Match 73.1%; Score 38; DB 20; Length 198;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
111:1:1
Db 157 c1svplvpc 165

RESULT 12

ABB57779 standard; Protein: 562 AA.

AC ABB57779:

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 129.

KW Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEXE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75.

DR N-PSDB; ABL01882.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Disclosure; SEQ ID NO 129; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
sequences (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 562 AA:

Query Match 73.1%; Score 38; DB 22; Length 562;
Best Local Similarity 55.6%; Pred. No. 2,1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 C1SVPLVPC 9
1:1:1
Db 9 cvkvpldpc 17

RESULT 13

AAY27020 standard; Protein: 3623 AA.

AC AAY27020;

DT 08-OCT-1999 (first entry)

DE Rat cubilin protein sequence.

KW Cubilin; epithelial glycoprotein receptor; vitamin B12; endocytosis;

KW toxicity; kidney; spleen; brain; liver; heart; thyroid; abortion; rat;

KW proteinuria; fetal malformation; fetal development; kidney damage.

XX Rattus sp.

FH Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "endoplasmatic import signal sequence"

FT Modified-site 95 /note= "mature protein"

FT Modified-site 428 /note= "potential N-glycosylation site"

FT Modified-site 491 /note= "potential N-glycosylation site"

FT Modified-site 711 /note= "potential N-glycosylation site"

FT Modified-site 781 /note= "potential N-glycosylation site"

FT Modified-site 857 /note= "potential N-glycosylation site"

FT Modified-site 957 /note= "potential N-glycosylation site"

FT Modified-site 984 /note= "potential N-glycosylation site"

FT Modified-site 1168 /note= "potential N-glycosylation site"

FT Modified-site 1285 /note= "potential N-glycosylation site"

FT Modified-site 1307 /note= "potential N-glycosylation site"

FT Modified-site 1319 /note= "potential N-glycosylation site"

FT Modified-site 1332 /note= "potential N-glycosylation site"

FT Modified-site 1500 /note= "potential N-glycosylation site"

FT Modified-site 1551 /note= "potential N-glycosylation site"

FT Modified-site 1646 /note= "potential N-glycosylation site"

FT Modified-site 1671 /note= "potential N-glycosylation site"

FT Modified-site 1802 /note= "potential N-glycosylation site"

FT Modified-site 1819 /note= "potential N-glycosylation site"

FT Modified-site 2085 /note= "potential N-glycosylation site"

FT Modified-site 2117 /note= "potential N-glycosylation site"

FT Modified-site /note= "potential N-glycosylation site"

FT	Modified-site	2274	"potential N-glycosylation site"
FT	/note=	2400	"potential N-glycosylation site"
FT	Modified-site	2531	"potential N-glycosylation site"
FT	/note=	2581	"potential N-glycosylation site"
FT	Modified-site	2610	"potential N-glycosylation site"
FT	/note=	2813	"potential N-glycosylation site"
FT	Modified-site	2875	"potential N-glycosylation site"
FT	/note=	2945	"potential N-glycosylation site"
FT	Modified-site	2989	"potential N-glycosylation site"
FT	/note=	3042	"potential N-glycosylation site"
FT	Modified-site	3106	"potential N-glycosylation site"
FT	/note=	3125	"potential N-glycosylation site"
FT	Modified-site	3165	"potential N-glycosylation site"
FT	/note=	3268	"potential N-glycosylation site"
FT	Modified-site	3283	"potential N-glycosylation site"
FT	/note=	3290	"potential N-glycosylation site"
FT	Modified-site	3357	"potential N-glycosylation site"
FT	/note=	3400	"potential N-glycosylation site"
FT	Modified-site	3430	"potential N-glycosylation site"
FT	/note=	3533	"potential N-glycosylation site"
FT	Modified-site	37	"potential N-glycosylation site"
FT	/note=	80	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	Misc-difference	108	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	/note=	148	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	Misc-difference	224	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	/note=	234	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	Misc-difference	235	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	/note=	241	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	Misc-difference	274	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	/note=	353	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	Misc-difference	358	"the His residue at this position is given as a Gly residue in the formal Seq ID listing"
FT	/note=		

[illegible]

Db 313 csqaplvpc 321

RESULT 14

AA000458
ID AAG00458 standard; Protein; 64 AA.

XX AAG00458;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 4539.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 990S-0122487.

XX (GSEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC00464.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 13; SEQ ID 4539; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 64 AA;

Query Match 71.2%; Score 37; DB 21; Length 64;

Best Local Similarity 77.8%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9

Db 5 c1svclpc 13

RESULT 15

AA012860
ID AAY12860 standard; Protein; 65 AA.

XX AAY12860;

XX 21-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO:450.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.

XX WO9906549-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01231.

XX 01-AUG-1997; 97US-0905279.

XX (GSEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153779/13.

XX N-PSDB; AAX51638.

PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from testis, ovary, uterus and spleen tissue
 XX Claim 34; Page 489; 522pp; English.

CC AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12861 to
 CC AAY12913, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 65 AA;

Query Match 71.2%; Score 37; DB 20; Length 65;

Best Local Similarity 77.8%; Pred. No. 36;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C1SVPLVPC 9

Db 5 c1svclpc 13

Search completed: June 17, 2002, 16:02:18
 Job time: 427 sec

